

Environment-by-PGS Interaction in the Classical Twin Design: An Application to Childhood Anxiety and Negative Affect

Susanne Bruins, Jouke-Jan Hottenga, Michael C. Neale, René Pool, Dorret I. Boomsma & Conor V. Dolan

To cite this article: Susanne Bruins, Jouke-Jan Hottenga, Michael C. Neale, René Pool, Dorret I. Boomsma & Conor V. Dolan (13 Jul 2023): Environment-by-PGS Interaction in the Classical Twin Design: An Application to Childhood Anxiety and Negative Affect, Multivariate Behavioral Research, DOI: [10.1080/00273171.2023.2228763](https://doi.org/10.1080/00273171.2023.2228763)

To link to this article: <https://doi.org/10.1080/00273171.2023.2228763>



© 2023 The Author(s). Published with license by Taylor & Francis Group, LLC



Published online: 13 Jul 2023.



Submit your article to this journal [↗](#)



Article views: 365



View related articles [↗](#)



View Crossmark data [↗](#)

Environment-by-PGS Interaction in the Classical Twin Design: An Application to Childhood Anxiety and Negative Affect

Susanne Bruins^{a,b} , Jouke-Jan Hottenga^a , Michael C. Neale^{a,c} , René Pool^{a,b} ,
Dorret I. Boomsma^{a,b,d} , and Conor V. Dolan^a 

^aDepartment of Biological Psychology, Vrije Universiteit; ^bAmsterdam Public Health research institute; ^cVirginia Institute for Psychiatric and Behavioral Genetics, Virginia Commonwealth University; ^dAmsterdam Reproduction and Development research institute

ABSTRACT

One type of genotype-environment interaction occurs when genetic effects on a phenotype are moderated by an environment; or when environmental effects on a phenotype are moderated by genes. Here we outline these types of genotype-environment interaction models, and propose a test of genotype-environment interaction based on the classical twin design, which includes observed genetic variables (polygenic scores: PGSs) that account for part of the genetic variance of the phenotype. We introduce environment-by-PGS interaction and the results of a simulation study to address statistical power and parameter recovery. Next, we apply the model to empirical data on anxiety and negative affect in children. The power to detect environment-by-PGS interaction depends on the heritability of the phenotype, and the strength of the PGS. The simulation results indicate that under realistic conditions of sample size, heritability and strength of the interaction, the environment-by-PGS model is a viable approach to detect genotype-environment interaction. In 7-year-old children, we defined two PGS based on the largest genetic association studies for 2 traits that are genetically correlated to childhood anxiety and negative affect, namely major depression (MDD) and intelligence (IQ). We find that common environmental influences on negative affect are amplified for children with a lower IQ-PGS.

KEYWORDS



Genotype-environment ($g \times e$) interaction; moderation; mono- and dizygotic twins; polygenic scores (pgs); genetic covariance structure modeling

Introduction

Genotype-environment interaction occurs when the effects of environmental exposure on a trait systematically depend on an individual's genotype, or when the effects of a genotype on a trait depend on an individual's environmental exposure (Eaves 1977; Boomsma and Martin 2002). On the substantive level, genotype-environment interaction is relevant to our understanding of development across multiple domains, including psychopathology (Hicks et al. 2009; Samek et al. 2015, 2017; D. Molenaar et al. 2016), personality (Boomsma et al. 1999), or substance use (Koopmans et al. 1999; Dick et al. 2007). On the statistical level, genotype-environment interaction is relevant to the correct interpretation of parameters in the modeling of data from twins, families, or genotyped individuals. Unmodeled genotype-environment interaction, if present, may result in biased parameter estimates (Purcell 2002; Verhulst et al. 2019), that depend on the source of the interaction. Here we often distinguish between environmental factors that

are common (C) to family members who share their environment, and environmental factors that are not shared between family members (E). An interaction of additive genotype (A) with common environmental factors ($A \times C$ interaction) presents as part of the additive genetic variance, and any $A \times E$ interaction presents as part of the unshared environmental variance.

The vast majority of genotype-environment interaction studies in behavioral science are conducted within a framework in which the effect of the genotype depends on a measured environmental exposure, as historically it was more feasible to measure an aspect of the environment than a person's genotype. In this framework, genotype-environment interaction can be tested by moderation models, where the environment moderates the influence of latent genetic and environmental factors on an outcome phenotype (Purcell 2002). For example, a religious upbringing diminishes genetic effects on disinhibition (Boomsma et al. 1999) as adolescent twins with a religious

CONTACT Susanne Bruins  s.bruins@vu.nl  Department of Biological Psychology, Vrije Universiteit, Amsterdam, The Netherlands

© 2023 The Author(s). Published with license by Taylor & Francis Group, LLC

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

upbringing report lower levels of disinhibition, and genetic effects on disinhibition are smaller in this group. Genetic effects on the development of adolescent externalizing disorders appear to be amplified in the face of environmental stress (Hicks et al. 2009; Samek et al. 2017). This moderating effect of environmental stress is important in adolescence, but wanes in young adulthood (Johnson et al. 2009; Samek et al. 2015, 2017).

The study of $G \times E$ interaction in genetically informative designs has been facilitated by the application of structural equation modeling (SEM; Boomsma and Molenaar 1986; Molenaar and Boomsma 1987). Boomsma and Molenaar showed how the covariance between measures and between family members could be decomposed into genetic and environmental sources of variance and covariance using the LISREL software (Jöreskog 1973; Boomsma and Molenaar 1986), and D. Molenaar and Dolan (2014) illustrated how methods and approaches as developed in psychometrics proved to be inspirational in the study of $G \times E$ interaction.

Methods that investigate genotype-environment interaction as a function of a measured (environmental) moderator are relatively intuitive and straightforward to implement, as outlined by Purcell (2002). Peter Molenaar demonstrated that genotype-environment interaction in the classical twin design can also be detected by nonlinear factor analysis (Molenaar and Boomsma, 1987), and by analysis of the distribution of factor scores (Molenaar et al. 1990, 1999) in cross-sectional and longitudinal structural equation models, without the need to measure the environmental exposure (or the genotype). Inspired by Peter Molenaar's work on factor scores, van der Sluis et al. (2012) and D. Molenaar et al. (2012a, 2013) proposed to study genotype-environment interaction by considering additive genetic factor (A) scores as moderators of environmental effects by marginal maximum likelihood estimation to condition on the latent A factor, rather than actually estimating genetic factor scores. This latent variable approach avoids the loss of information incurred by a two-step analysis of deriving and analyzing factor scores, when the scores may not be identically distributed.

In this article, we turn to consider the SEM implementations of genotype-environment interaction, where the effects of latent variables (genotypes and environments) are moderated by measured genotypes at the individual level. With high throughput genotyping technologies, it is feasible to measure genetic variants in large samples. These usually are single

nucleotide polymorphisms (SNPs), a single nucleotide in the genome sequence that varies between people. SNPs that are associated with a phenotype of interest can be combined into polygenic scores (PGS) which, as an observed variable, provides a (partial) handle on the otherwise latent genetic variable underlying a phenotype of interest (Winkler et al. 2014; Tam et al. 2019; Choi et al. 2020).

Here we consider a model that incorporates PGS in structural equation models for data collected in mono- and dizygotic twin pairs to test for genotype-environment interaction. In the approach we take in this paper a PGS, an observed genetic variable, functions as the moderator. This allows us to test the hypothesis that the effect of the environment on the phenotype of interest depends on an individual's genotype. Because we apply the model in the context of twin data, we test both $A \times C$ and $A \times E$ interaction. The PGS may reflect the genetic contribution to the phenotype that is analyzed, e.g., a PGS for major depression disorder (MDD) based on a genome-wide association study (GWAS) for MDD, or may reflect the genetic contribution to a different phenotype that is relevant to the phenotype of interest. Most human traits are polygenic and multiple traits tend to be correlated for genetic reasons (Brainstorm Consortium et al. 2018). Therefore, PGSs for one trait tend to correlate with a variety of other traits.

This paper is organized as follows. First, we introduce classical twin design and the concept of PGSs. We refrain from a detailed description of PGS estimation, as excellent resources are available (Choi et al. 2020). We then describe a test of genotype-environment interaction in which the PGS functions as a genetic moderator. We present a simulation study, in which we address parameter recovery and statistical power of the Environment-by-PGS interaction model, and end with an empirical study in which we apply this model to data on childhood anxiety and negative affect.

The twin classical design

The classical twin design (CTD) is applied to decompose the phenotypic covariance matrix of twins into genetic and environmental covariance matrices (Bruins et al. 2022a). The basis for the decomposition is the difference in genetic relatedness between two types of twins: monozygotic (MZ) twins arise from a single fertilized egg and therefore share 100% of their DNA, except for rare somatic mutations that arise in

one member of a twin pair, but not in the other. Dizygotic (DZ) twins share on average 50% of their alleles at segregating loci. Both types of twins may share environmental influences, including prenatal exposures, which may contribute to their phenotypic resemblance. All environmental influences that contribute to phenotypic differences within pairs are referred to as unshared (unique) environments. We typically distinguish four sources of variance, namely additive genetic (A) variance (a^2), non-additive or dominance genetic (D) variance (d^2), common environmental (C) variance (c^2), and unshared (E) environmental variance (e^2) that typically also includes measurement error variance. The model including all four variance components is not identified with only data from MZ and DZ twin pairs. Here, we proceed with an ACE model, where the phenotypic variance then equals $\sigma_Y^2 = a^2 + c^2 + e^2$, when assuming no covariance or interaction among variance components (Eaves 1977). Figure 1 displays a path diagram of the univariate ACE twin model. The expected MZ and DZ covariance matrices are:

$$\begin{aligned} \Sigma_{MZ}^2 &= \begin{pmatrix} \sigma_{MZ}^2 & \sigma_{MZ1, MZ2} \\ \sigma_{MZ1, MZ2} & \sigma_{MZ}^2 \end{pmatrix} \\ &= \begin{pmatrix} a^2 + c^2 + e^2 & a^2 + c^2 \\ a^2 + c^2 & a^2 + c^2 + e^2 \end{pmatrix} \end{aligned} \quad (1)$$

$$\begin{aligned} \Sigma_{DZ}^2 &= \begin{pmatrix} \sigma_{DZ}^2 & \sigma_{DZ1, DZ2} \\ \sigma_{DZ1, DZ2} & \sigma_{DZ}^2 \end{pmatrix} \\ &= \begin{pmatrix} a^2 + c^2 + e^2 & 0.5a^2 + c^2 \\ 0.5a^2 + c^2 & a^2 + c^2 + e^2 \end{pmatrix} \end{aligned} \quad (2)$$

The contributions of A to the phenotypic covariance between twins correspond to the expected proportions of shared alleles in the MZ (1) and DZ twin pairs (0.5).

Observed genetic variables: Polygenic scores

At a locus with 2 alleles, denoted B and b, we can observe 3 genotypes: bb, bB or Bb, and BB, which are coded 0 (bb), 1 (Bb or bB), and 2 (BB). An individual's PGS is calculated as the weighted linear combination of these values from all SNPs that contribute to a phenotype (Choi et al. 2020; Wray et al. 2021). The weights are based on effects estimated in genome-wide association studies (see, e.g., Vilhjálmsson et al. 2015). Suppose that K SNPs contribute to the genetic variance of a trait. We express an individual's (i) phenotype (Y) as:

$$Y_i = b_0 + \sum_{k=1}^K b_k \text{SNP}_{ki} + cC_i + eE_i, \quad (3)$$

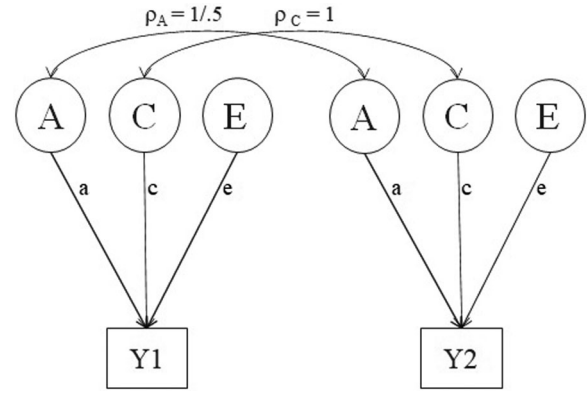


Figure 1. Path diagram of the ACE twin model. Y indicates the measured phenotype (Y1 and Y2 are the phenotypes of twin 1 and twin 2). Squares denote observed variables; circles denote latent variables. ρ_A denotes the degree of genetic similarity (correlation) between members of a twin pair (1 for MZ and 0.5 for DZ twins); A, C, and E are uncorrelated and represent the genetic, shared environmental, and unshared environmental factors; a, c, and e are the additive genetic, shared environmental, and unshared environmental path coefficients respectively. The variance of the latent variables is 1.

where C and E are standardized latent variables, SNP (a measured locus with three distinct genotypes) is coded 0, 1, or 2, and b_k is the regression coefficient of the kth SNP. The phenotypic variance is expressed as

$$\sigma_Y^2 = \sum_{k=1}^K b_k^2 \sigma_{\text{SNP}k}^2 + c^2 + e^2 \quad (4)$$

In practice, we will have a subset comprising P of the K SNPs, that are used in calculating the PGS, and with $L = K - P$ unmeasured SNPs contributing to the phenotypic variance. Then

$$\begin{aligned} \sigma_Y^2 &= \sum_{p=1}^P b_p^2 \sigma_{\text{SNP}p}^2 + \sum_{l=1}^L b_l^2 \sigma_{\text{SNP}l}^2 + c^2 + e^2 \\ &= a_p^2 + a_l^2 + c^2 + e^2 \end{aligned} \quad (5)$$

where $a_p^2 = \sum_{p=1}^P b_p^2 \sigma_{\text{SNP}p}^2$ is the genetic variance of Y explained by the PGSs, and $a_l^2 = \sum_{l=1}^L b_l^2 \sigma_{\text{SNP}l}^2$ is the residual genetic variance of Y.

Environment-by-PGS interaction

We propose to model the $\text{PGS} \times C$ and $\text{PGS} \times E$ interaction, as depicted in Figure 2. The moderation (or interaction) parameters, denoted b_c and b_e , accommodate the dependency of the C and E effects on the PGS. We test this moderation by specifying the effects of C and E on the phenotype as a function of the PGS: $C = c_0 + b_c \times \text{PGS}$ and $E = e_0 + b_e \times \text{PGS}$. In the absence of environment-by-PGS interaction, the parameters c_0 and e_0 are equal to the parameters c and e in Figure 1. In Figure 2, the proportion of

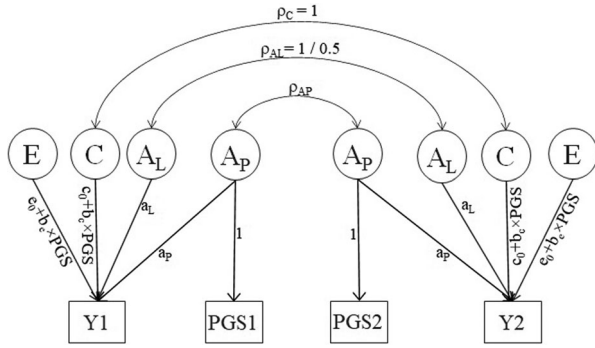


Figure 2. Path diagram for environment-by-PGS interaction model. Y denotes the phenotypes of twin 1 and twin 2. PGS1 and PGS2 are polygenic scores of twins 1 and 2. Squares denote observed variables and circles denote latent variables. A_P and the PGS are identical (one-to-one relationship), and A_L , C , and E are the latent variables with effects a_L , c_0 , and e_0 , respectively. The parameter a_P denotes the main effect of the PGS on the phenotype; b_c and b_e are the interaction parameters. The variances of all latent variables are 1. The parameter ρ_{AL} denotes the correlation between the additive genetic variables (1 for MZ twins and 0.5 for DZ twins), ρ_{AK} is the correlation between the PGSs.

genetic variance captured by the PGS is $R_A^2 = a_P^2 / (a_L^2 + a_P^2)$. The variance-covariance matrices are

$$\begin{aligned} \Sigma_{MZ} &= \begin{pmatrix} \sigma_{MZ}^2 & \sigma_{MZ1, MZ2} \\ \sigma_{MZ1, MZ2} & \sigma_{MZ}^2 \end{pmatrix} \\ &= \begin{pmatrix} a_L^2 + a_P^2 + (c_0 + b_c \times \text{PGS})^2 + (e_0 + b_e \times \text{PGS})^2 & a_L^2 + a_P^2 + (c_0 + b_c \times \text{PGS})^2 \\ a_L^2 + a_P^2 + (c_0 + b_c \times \text{PGS})^2 & a_L^2 + a_P^2 + (c_0 + b_c \times \text{PGS})^2 + (e_0 + b_e \times \text{PGS})^2 \end{pmatrix} \end{aligned} \quad (6)$$

$$\begin{aligned} \Sigma_{DZ} &= \begin{pmatrix} \sigma_{DZ}^2 & \sigma_{DZ1, DZ2} \\ \sigma_{DZ1, DZ2} & \sigma_{DZ}^2 \end{pmatrix} \\ &= \begin{pmatrix} a_L^2 + a_P^2 + (c_0 + b_c \times \text{PGS})^2 + (e_0 + b_e \times \text{PGS})^2 & 0.5a_L^2 + 0.5a_P^2 + (c_0 + b_c \times \text{PGS})^2 \\ 0.5a_L^2 + 0.5a_P^2 + (c_0 + b_c \times \text{PGS})^2 & a_L^2 + a_P^2 + (c_0 + b_c \times \text{PGS})^2 + (e_0 + b_e \times \text{PGS})^2 \end{pmatrix} \end{aligned} \quad (7)$$

We investigate $\text{PGS} \times C$ and $\text{PGS} \times E$ interaction by testing $b_c = 0$ and/or $b_e = 0$. Because $R_A^2 < 1$, the estimated interaction parameters (b_c , b_e) underestimate the true interaction parameter values. Considering $A \times C$ interaction, let \mathbf{A} denote the total additive genetic variable (i.e., $\mathbf{A} = \mathbf{A}_P + \mathbf{A}_L$, where \mathbf{A}_P is PGS), let β_c denote the true interaction parameter value, and let b_c denote the estimated interaction

parameter value. The effect of C and its interaction with A on the genotype is $c_0 + \beta_c \times \mathbf{A}$, but we estimate $\hat{c}_0 + b_c \times \mathbf{A}_P = \hat{c}_0 + b_c \times \sqrt{R_A^2} \times \mathbf{A}$. Thus, we underestimate the interaction parameter value by a factor of $\sqrt{R_A^2}$, such that $b_c = \beta_c \times \sqrt{R_A^2} = \beta_c \times R_A$, and we can find the true interaction parameter by dividing the estimate by R_A . The same applies to be and β_e :

$$\beta_c = \frac{b_c}{R_A} \text{ and } \beta_e = \frac{b_e}{R_A} \quad (8)$$

Note that this correction based on R_A is only applicable if the PGS is based on a GWAS of the phenotype of interest. If the PGS is based on a phenotype different from the phenotype of interest, R_A will be underestimated, and the resulting corrected interaction parameters will be inflated.

Simulation study

We explored the power to detect genotype-environment interaction in a simulation study and assume the following: Random mating of parents of twins, for the trait that is analyzed MZ twins are not treated dif-

ferently than DZ twins (equal environment assumption), no genotype-environment correlation, and the phenotypic data conditional on the genetic variable follow a multivariate normal distribution.

We considered various scenarios, in which we varied the genetic and environmental variance (a^2 , c_0^2 , e_0^2), the proportion of genetic variance explained by the PGSs (R_A^2), and the sample composition (MZ:DZ

Table 1. Glossary of terminology and notation.

Glossary	Notation
Genotype-environment correlation. Correlation between genes and environment that arises when exposure to certain environments is partially influenced by an individual's genotype.	<i>Y</i> . Observed phenotypic variable.
Genotype-environment interaction. Interaction between genotype and environment that can occur when the effects of the environment on a phenotype depend on genotype, or when the effects of the genotype on a phenotype depend on environment.	PGS . Measured PGS variable.
Single nucleotide polymorphism (SNP). A common type of genetic variation that occurs when a single nucleotide (Adenine, Thymine, Cytosine, or Guanine) in the DNA sequence is altered.	<i>A_L</i> . Latent additive genetic variable.
Genome-wide association study (GWAS). A study in which the associations between many genetic variants (often SNPs) and a phenotype are estimated.	<i>a_L</i> . Additive genetic path coefficient.
Polygenic score (PGS). Weighted linear combination of the SNPs that affect a phenotype. Also known as Polygenic Risk Score (PRS) or Polygenic Index (PGI).	<i>a_p</i> . Latent variable representing the PGS.
Additive genetic effects. Contribution of individual genes or alleles to a phenotype in a linear and additive manner.	<i>a_p</i> . PGS path coefficient, denoting the effect of the PGS on phenotype <i>Y</i> .
Dominance genetic effects. Non-additive genetic effects stemming from intra-locus interactions.	C . Common environmental variable.
Common environmental effects. Environmental effects that are shared between members of a twin pair.	<i>c_p</i> . Common environmental path coefficient.
Unique environmental effects. Environmental effects that are not shared between members of a twin pair, typically also includes measurement error.	<i>b_c</i> . C-by-PGS coefficient, denoting the interaction between the PGS and the common environment.
	E . Unique environmental variable.
	<i>e_p</i> . Unique environmental path coefficient.
	<i>b_e</i> . E-by-PGS coefficient, denoting the interaction between the PGS and the unique environment.
	<i>ρ_{AL}</i> . Additive genetic correlation between members of a twin pair. This corresponds to the proportion of shared genetic material (i.e., 1 for MZ twins and .50 for DZ twins).
	<i>ρ_{AP}</i> . Correlation between PGSs of a twin pair. Its value is expected to be 1 in MZ twins and .50, on average, in DZ twins.
	<i>δ_c</i> . Effect size of C-by-PGS interaction. Indicates the change in C variance with every standard deviation increase in the PGS.
	<i>δ_e</i> . Effect size of E-by-PGS interaction. Indicates the change in E variance with every standard deviation increase in the PGS. When the phenotype is scaled to have unit variance, <i>δ_c</i> and <i>δ_e</i> correspond to the proportion of phenotypic variance that is explained by the C- or E-by-PGS interaction.
	<i>R_A²</i> . Proportion of additive genetic variance captured by the PGS.

Note. Parameters in italics are freely estimated in the environment-by-PGS interaction model.

twin ratio). The interaction effect size (δ_c , δ_e) indicates the change in the environmental variance (c^2 , e^2) for each standard deviation increase in the PGS, i.e., the phenotypic variance that is explained by the interaction parameter ($\delta_c = b_c^2 + 2b_c c_0$ and $\delta_e = b_e^2 + 2b_e e_0$). The effect size was $\delta_c = \delta_e = 0.10$ over all simulations, and we simulated data for 4000 twin pairs. The power to detect environment-by-PGS interaction was calculated as the power to reject the model in which the non-zero interaction parameters (b_c , b_e) were fixed to 0, over 1,000 replications, given $\alpha = 0.05$, for separate (1 df) tests and an omnibus (2 df) test of the interaction parameters. An overview of simulation parameters is presented in Table 2. The choice of the parameter values is based in part on empirical parameter values from an earlier study where we found that IQ moderated genetic and environmental influences on psychopathology in children (Bruins et al. 2022b).

To determine whether the power to detect an interaction effect depended on sample composition (i.e., the MZ: DZ ratio), we considered the following scenarios: an equal number of MZ and DZ pairs (MZ: DZ = 1: 1); the ratio in the population with more DZ than MZ pairs (MZ: DZ = 1/3: 2/3); and an overrepresentation of MZ pairs relative to DZ pairs, as often found in adult twin studies (MZ: DZ = 2/3: 1/3).

In addition, we determined the false positive rate by simulating data without an interaction effect and calculated the proportion of simulations in which the

test of the interaction parameters had a p-value smaller than .05. To explore violation of the conditional multivariate distribution, we determined the false-positive rate on left-censored data. Specifically, we introduced a floor effect by assigning a score corresponding to the quantile associated with the probability of 0.15 to the lowest 15% of the phenotypic data, and again calculated the proportion of false positive results. We considered left-censoring because many measures of psychopathology display floor effects in population-representative samples.

All analyses were conducted in R 3.6.0 (R Core Team 2018) in the R-packages doSNOW (Analytics and Weston 2014), foreach (Calaway et al. 2015), MASS (Venables and Ripley 2013) for the simulation study, and OpenMx 2.17.1 for genetic model fitting (Neale et al. 2016). R scripts are available at Open Science Framework (OSF), via doi: 10.17605/OSF.IO/GB7WQ

Results of the simulation study

The power to detect the environment-by-PGS interaction effect depended on the A, C, and E variance, and increased with a greater proportion of genetic variance explained by the PGSs (R_A^2). We present the results of the simulation study in Table 3. In general, power to detect E-by-PGS interaction was greater than the power to detect C-by-PGS interaction, and the omnibus (2 df) test of b_c and b_e had greater power

Table 2. Parameter values in the simulation study.

Parameters settings			
Parameter	Values		
a^2	.35	.55	
c_0^2	.10	.35	
e_0^2	.35	.50	
R_A^2	.10	.20	
Parameters in the power simulations			
δ_c	.10		
δ_e	.10		
MZ/DZ ratio	$2/3 : 1/3$	$1/3 : 2/3$	1 : 1
Parameters in the false positive rate simulations			
δ_c	0		
δ_e	0		
MZ/DZ ratio	1 : 1		
Proportion of floor values	0	.15	

a^2 = total genetic variance; c_0^2 = shared environmental variance; e_0^2 = unshared environmental variance; R_A^2 = proportion of genetic variance explained by PGS; δ_c = C-by-PGS interaction effect size; δ_e = E-by-PGS interaction effect size. The total number of MZ + DZ twin pairs was 4000.

to detect environment-by-PGS interaction than the 1 df tests of individual interaction parameters (especially when the power for the individual tests was relatively modest, e.g., in simulations 2 and 8). Power also increases when the relative effect size increases. That is, the greater the ratio of parameters (i.e., b_c/c_0 and b_e/e_0) the greater the power to detect an environment-by-PGS interaction effect. Power was slightly lower when the MZ:DZ ratio was unbalanced (i.e., other than 1:1; see [Supplementary Table 1](#)). [Figure 3](#) displays examples of conditional variance and power (based on simulation 1), and parameter bias in the interaction parameters (stemming from $R_A^2 < 1$).

Given multivariate normal phenotypic data (conditional on the PGS), the average false-positive rate is approximately equal to the alpha of 0.05 (fluctuating from 0.03 to 0.07 across simulations). However, when we introduced a floor effect in the data by collapsing the lowest 15% of the phenotypic scores into a single minimum score, the false positive rate ranged from 0.5 to 1 across simulations (see [Supplementary Table 2](#)).

Empirical application to childhood anxiety and negative affect

We analyzed data on anxiety and affective problems from seven-year-old twins. These are heritable traits ([Middeldorp et al. 2005](#); [Wesseldijk et al. 2017](#)). [Bruins et al. \(2022a, 2022b\)](#) found that genetic and/or environmental effects on anxiety and affect at age 7 were moderated by a total IQ score, as assessed by psychometric intelligence tests for children. Here, we tested whether environmental effects on anxiety and affect are moderated by a PGS of IQ. The discovery GWA study ([Savage et al. 2018](#)) published the effect

sizes for SNPs for IQ in adults and children, which we used to calculate PGSs for the twins in our sample. The second series of tests is considered a PGS based on the currently largest discovery GWAS for Major Depressive Disorder ([Wray et al. 2018](#)). These discovery GWASs were mainly carried out in adult participants, but there is substantial genetic stability from early childhood to adulthood for IQ as well as for symptoms of anxiety and depression (e.g., [Nivard et al. 2015](#)), indicating that largely the same genetic variants are expressed in children and adults.

Participants

Genotype and phenotype data were available for 1391 MZ and 1185 DZ twin pairs who participate in the Young Netherlands Twin Register (YNTR). Participants of the YNTR are recruited at birth, and their parents complete surveys on their development until the twins reach the age of 12 years. In this study, we analyzed data on anxiety and affective problems, as reported by mothers on their 7-year-old children ([Van Beijsterveldt et al. 2013](#)).

Measurements

Anxiety and negative affect were assessed by the DSM-oriented scales of the Child Behavior Check List (CBCL; [Achenbach et al. 2003](#)). The CBCL includes 112 items concerning specific problem behaviors that are rated on a scale from 0–2, where 0 = not true; 1 = somewhat or sometimes true; and 2 = very true or often true. The affect scale (“affective problems”) consisted of 13 items, and the anxiety scale (“anxiety problems”) consisted of 6 items. Both scales have good test-retest reliability over 8 days ($r = 0.84$ for affective problems and $r = 0.80$ for anxiety problems), and acceptable to good internal consistency ($\alpha = 0.82$ for affective problems and $\alpha = 0.72$ for anxiety problems; [Achenbach 2001](#)). Total scale scores were constructed with item response theory, using a graded response model ([Samejima 1997](#)). The distribution of the scores is characterized by left censoring (a floor effect), caused by overuse of the “not true” category. We return to the effect of this censoring on the results of our analyses below.

The PGSs were based on the discovery GWA meta-analysis of ([Savage et al. 2018](#)) for intelligence, and the GWAS of ([Wray et al. 2018](#)) for major depressive disorder. We detail our procedure for PGS construction in the [supplementary methods](#). We scaled the PGSs to have a mean of 0 and a variance of 1, and we

Table 3. Results of simulation study, averaged over 1000 replications.

Parameter settings	Estimates										Relative effect size		Power			n .80		
	R_A^2	a^2	c_0^2	e_0^2	β_c	β_e	b_c	b_e	b_c^*	b_e^*	b_c/c_0	b_e/e_0	b_c	b_e	om	b_c	b_e	om
1	.1	.55	.10	.35	.131	.079	.044	.025	.143	.082	.14	.04	.45	.90	.98	12982	3300	2306
2	.1	.55	.10	.50	.131	.067	.050	.021	.161	.068	.16	.03	.45	.68	.92	13025	6433	3341
3	.1	.55	.35	.35	.079	.079	.030	.025	.094	.079	.05	.04	.52	.88	.97	10121	3565	2450
4	.1	.55	.35	.50	.079	.067	.023	.022	.074	.07	.04	.03	.33	.70	.86	24257	5994	4196
5	.1	.35	.10	.35	.131	.079	.042	.025	.134	.081	.13	.04	.58	.91	.99	8379	3119	1750
6	.1	.35	.10	.50	.131	.067	.043	.021	.139	.067	.14	.03	.51	.73	.97	10694	5513	2537
7	.1	.35	.35	.35	.079	.079	.028	.025	.089	.078	.05	.04	.55	.89	.99	9137	3331	2132
8	.1	.35	.35	.50	.079	.067	.027	.020	.084	.063	.05	.03	.44	.65	.90	13651	6832	3532
9	.2	.55	.10	.35	.131	.079	.063	.035	.143	.079	.20	.06	.73	.99	1	5489	1716	1084
10	.2	.55	.10	.50	.131	.067	.063	.031	.143	.071	.20	.04	.64	.94	1	7191	2796	1522
11	.2	.55	.35	.35	.079	.079	.036	.035	.080	.079	.06	.06	.64	.99	1	7183	1660	1252
12	.2	.55	.35	.50	.079	.067	.035	.029	.079	.065	.06	.04	.54	.89	.99	9420	3421	2110
13	.2	.35	.10	.35	.131	.079	.058	.036	.130	.080	.18	.06	.84	.99	1	3974	1508	831
14	.2	.35	.10	.50	.131	.067	.056	.030	.128	.068	.18	.04	.69	.93	1	6197	2879	1333
15	.2	.35	.35	.35	.079	.079	.037	.035	.083	.079	.06	.06	.74	.99	1	5333	1610	1074
16	.2	.35	.35	.50	.079	.067	.038	.030	.086	.067	.06	.04	.69	.92	1	6149	2994	1584

Note. $n_{MZ} = n_{DZ} = 2000$ (i.e., 4000 twin pairs, 8000 individuals). R_A^2 is the proportion of genetic variance explained by the PGSs. β_c and β_e are the true interaction parameters, b_c and b_e are the estimated interaction parameters, and b_c^* and b_e^* are the estimated parameters corrected for R_A , om stands for omnibus (2 df) test, n 0.80 is the number of twin pairs (MZ + DZ) required to detect an interaction effect with a power of 0.80 (given $\alpha = 0.05$).

reverse-coded the IQ PGS such that it correlated positively with negative affect and anxiety.

The best-fitting twin model for anxiety includes the influence of A and E, while for negative affect we also find a contribution of C (Bruins et al. 2022b). Therefore, we tested only E-by-PGS interaction for anxiety, and both E-by-PGS and C-by-PGS interaction for negative affect. As demonstrated in the simulation results above, censoring can lead to false positive interaction parameters. When we detected environment-by-PGS interaction, we took the effect of censoring into account by re-analyzing the data while explicitly modeling the left-censored phenotypic distribution in OpenMx (see for example, de Zeeuw et al. 2019). For all analyses, we used a significance threshold of $\alpha = 0.05$.

Results application to empirical data

The PGSs for IQ and MDD each captured 0.3% of the additive genetic variance of anxiety, and 0.4% of the additive genetic variance of negative affect. The correlation between these PGSs equaled 0.135. Results for negative affect suggested some interesting interaction effects, both for the IQ-PGS and the MDD-PGS. There was C-by-PGS interaction, where shared environmental variance for negative affect changed with a higher PGS for negative affect or lower IQ, such that most shared environmental variance of negative affect can be observed in the tails of the PGS distribution. For negative affect, there was also evidence for E-by-PGS interaction, where unshared environmental variance increased with a lower PGS for IQ. The E-by-PGS interaction coefficient for the MDD PGS was

not significant ($p = 0.063$). For anxiety, we observed no environment-by-PGS interaction.

In the simulation study, we demonstrated that the false-positive rate drastically increases when there is a floor effect in the data. In our empirical data, 15% of the subjects have the minimum score on negative affect (see Figure 4). To assess how this floor effect influences the results, we re-analyzed the negative affect data, while explicitly modeling the left-censored phenotypic distribution. We still found that there was C-by-IQ-PGS interaction after correcting for censoring, but C-by-MDD-PGS interaction and E-by-PGS interaction disappeared. The model fitting results are in Table 4. The conditional variance components are shown in Figure 5.

Given that the PGSs explained less than 1% of the genetic variance of anxiety and negative affect, and that the sample size was small compared to the simulated sample size, the power to detect a small effect was low, and the obtained estimates are rather imprecise. When we correct the interaction parameters for the proportion of genetic variance explained by the PGSs, the resulting corrected estimates are implausibly large. This is in line with our simulations, in which underpowered studies yield imprecise estimates. When corrected for R_A , these corrected estimates may not reflect the true effect.

Discussion

Our aim was to present the environment-by-PGS interaction model and to discuss its implementation in combination with the classical twin design. We looked at the power to detect environment-by-PGS interaction for two different types of environmental

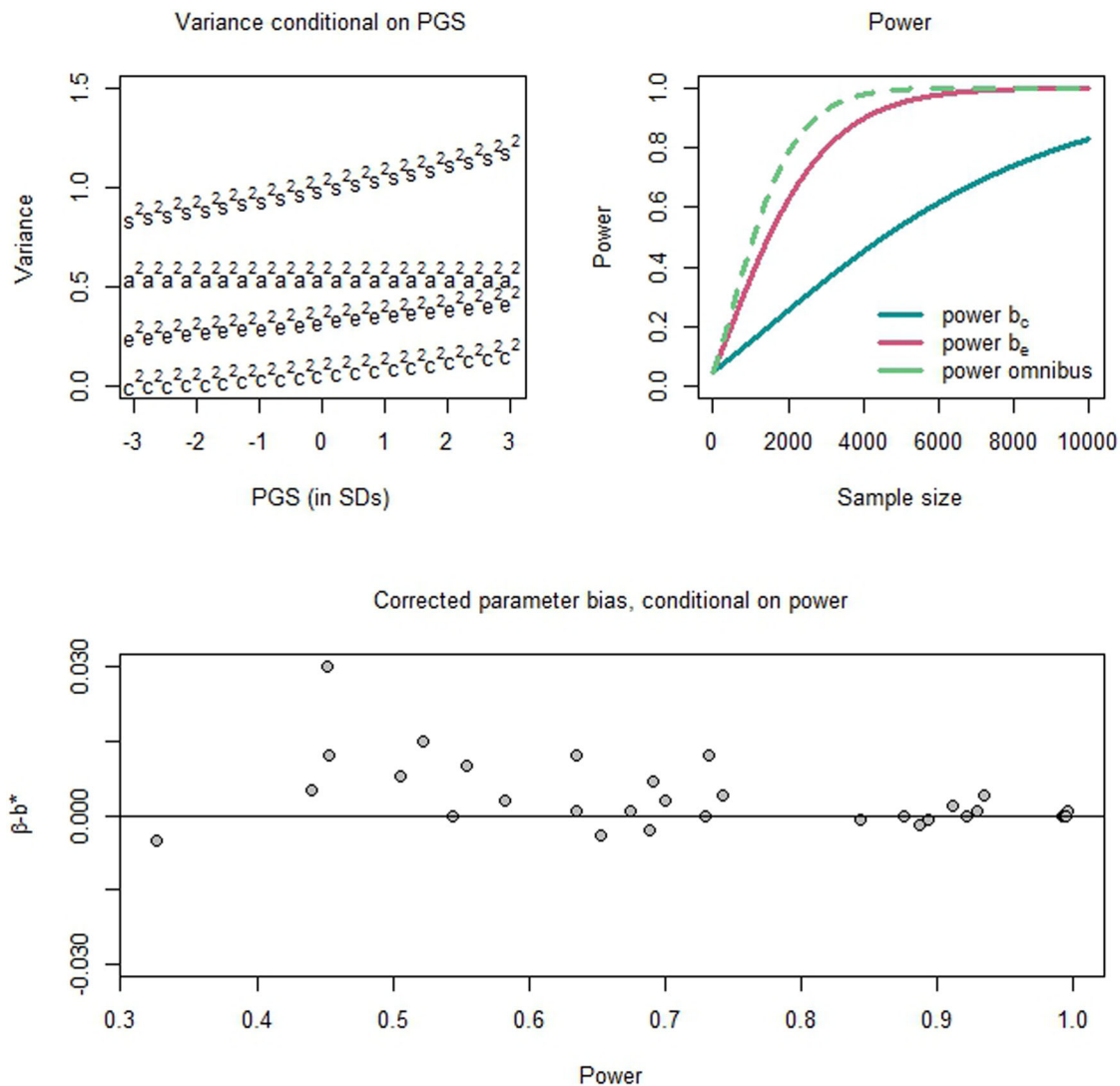


Figure 3. Results from the simulation study. **Top-left:** Genetic and environmental variance conditional on the PGS, from simulation 1. X-axis is the PGS (in SDs), y-axis is variance. a^2 = total additive genetic variance ($a_L^2 + a_P^2$), c^2 = common environmental variance, e^2 = unshared environmental variance, and s^2 = total phenotypic variance. **Top-right:** The power to detect a non-zero b_c , b_e , and b_{omnibus} parameter, in simulation 1. X-axis is number of twin pairs (MZ + DZ), y-axis is statistical power. **Bottom:** The discrepancy between the true interaction parameter (β as simulated) and the corrected estimate (i.e., b/R_A), conditional on the simulation power. Power is given on the X-axis, and the bias is shown on the y-axis ($\beta - b^*$).

influences, those that create a resemblance between children from the same household (common environment) and those that are unique to a person (unique environment). We found that the detection of genotype-environment interaction depends on the size of the A, C, and E effects, and on the genetic variance explained by PGSs. We found that the power to detect E-by-PGS generally was greater than the power to detect C-by-PGS interaction, and a 2 df omnibus test of both C- and E-by-PGS interaction had the greatest statistical power.

The finding that the power to detect E-by-PGS interaction is generally greater than the power to detect C-by-PGS interaction is in line with previous results of Molenaar et al (D. Molenaar et al. 2012; D.

Molenaar and Dolan 2014). When there is no particular interest in the type of interaction (E-by-PGS or C-by-PGS), we recommend testing for environment-by-PGS interaction with a 2 df omnibus test. The power to detect environment-by-PGS interaction also depends on the proportion of additive genetic variance (R_A^2) that a PGS explains. As of current, this proportion is typically small, but with ongoing progress in GWAS meta-analyses, the R_A^2 is likely to increase. We noted that the estimates of the interaction parameters (b_c and b_e) depend on the genetic variance explained by the PGS (R_A^2). However, as we demonstrated, it is simple to correct these parameter estimates, given that the parameters are estimated with reasonable accuracy.

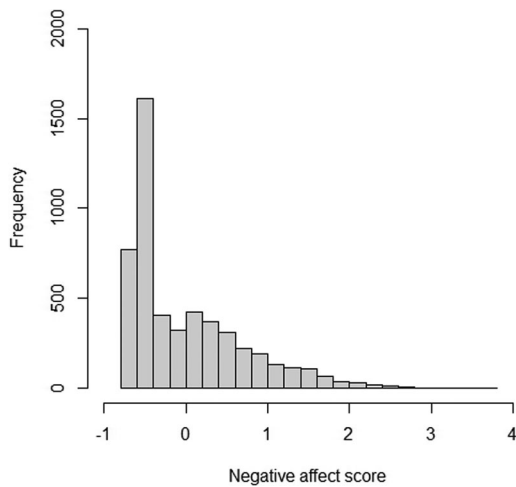


Figure 4. Histogram of negative affect scores.

We applied the environment-by-PGS interaction model to empirical data on childhood anxiety and negative affect, recognizing that although the study included 5,152 children with genotype and phenotype data, the power calculations from our simulations and the highly skewed phenotypic data suggest that caution is required. We considered two PGSs. Based on earlier work in which phenotypic IQ was tested as a moderator, we considered an IQ PGS, and the other PGS was based on a large GWAS in adults for major depressive disorder. Although these two PGSs explained less than 1% of the genetic variance, we detected C-by-PGS interaction for negative affect, where a PGS for lower IQ was associated with more shared environmental variance contributing to individual differences in negative affect. We accommodated the floor effect in the phenotypic data by modeling a censored distribution. However, our data did not follow a censored normal distribution, and our empirical results may yet be influenced by distributional issues.

Most genotype-environment interaction studies in behavioral science have focused on how genetic effects are expressed differentially depending on the environment (environmental moderation of genetic effects; Hagenbeek et al. 2022; Willoughby et al. 2023). The environment-by-PGS interaction method provides an explicit test of the hypothesis that individuals may have a differential sensitivity to environmental circumstances depending on their genotype. One advantage of this method is that a prior hypothesis concerning which specific environments have differential effects across levels of genetic sensitivity is not required. As such, studies using this method can serve as a first step in determining whether environmental

effects differ over levels of genetic sensitivity, and if so, follow-up studies can focus on identifying specific environmental circumstances.

The present method is a contribution to the growing body of methods that exploit measured genetic information. While these strategies have gained popularity for the study of genotype-environment correlation (Dolan et al. 2021), its application in the study of genotype-environment interaction has been more limited (Hagenbeek et al. 2022). The incorporation of PGSs in the modeling of family data has already allowed for the development of new and more powerful tests of genotype-environment correlation. PGSs in family members have been leveraged to estimate the covariance between shared environment and genotype in twins and full sibs (Selzam et al. 2019; Balbona et al. 2021; Dolan et al. 2021) and in parents and offspring. (Kong et al. 2018; Bates et al. 2018; de Zeeuw et al. 2019).

Rathouz et al (2008) have presented a careful analysis of Purcell's (2002) original bivariate moderation model. They demonstrated that a non-linear relationship between the moderator M and the phenotype may give rise to spurious A-by- M interaction. This spurious A-by- M interaction specifically concerns the moderation of the common A effects on the phenotype. By common A we mean the A component that is common to the moderator and the phenotype in the standard Cholesky representation of the Purcell model. However, in the present model, the effects of the PGS are not subject to moderation, and interaction is limited to phenotype-specific C and E effects. Hence, these concerns do not apply to the environment-by-PGS interaction model.

The tests of environment-by-PGS interaction, like most tests of interaction, are scale-dependent (Wagenmakers et al. 2012). That is, genotype-environment interaction estimates depend on the scaling of the phenotype, and positive results may be a function of "bad" scaling (inconsistent measurement precision across the scale) rather than true interaction (Eaves 2006). We considered a censored model, to accommodate the evident floor effects. While this is relatively simple to implement, we acknowledge there are other possibly more appropriate methods that one may consider to reduce the probability of false positive interaction tests. A potential solution to overcome scaling issues is to introduce a psychometric model to link phenotypic indicators (e.g., items) to the phenotypic latent variable. The test of interaction can then be conducted at the level of the latent variable (van den Berg et al. 2007; Molenaar and Dolan 2014). An

Table 4. Model fitting results of environment-by-PGS analyses for anxiety and negative affect.

	Phenotype	Test	-2LL	df	χ^2	Δ df	<i>p</i>	AIC	b (b*)	
IQ PGS	Anxiety		21236.79	8906				3424.79		
		b_e	21238.74	8907	1.94	1	.164	3424.74	-0.01	
	Neg. affect		20332.49	8904				2524.49		
		b_c	20337.03	8905	4.54	1	.033	2527.03	.11 (1.92)	
	Neg. affect censored dist.		b_e	20336.36	8905	3.87	1	.049	2526.36	.01 (0.24)
				21158.46	8904				3350.46	
		b_c	21163.61	8905	5.16	1	.023	3353.61	.13 (2.63)	
		b_e	21161.32	8905	2.87	1	.090	3351.32	.01	
MDD PGS	Anxiety		21324.64	8906				3512.64		
		b_e	21324.87	8907	0.23	1	.631	3510.87	-0.004	
	Neg. affect		20406.68	8904				2598.68		
		b_c	20413.73	8905	7.05	1	.008	2603.73	.11 (1.29)	
	Neg. affect censored dist.		b_e	20410.13	8905	3.45	1	.063	2600.13	.01
				21240.08	8904				3432.08	
		b_c	21243.68	8905	3.60	1	.058	3433.68	.11	
		b_e	21241.44	8905	1.35	1	.245	3431.44	.01	

-2LL is the -2 loglikelihood, df is degrees of freedom, b is the interaction parameter, and b* is the interaction parameter corrected by R_A . Neg. Affect censored dist. refers to the analyses in which we modeled a left-censored distribution.

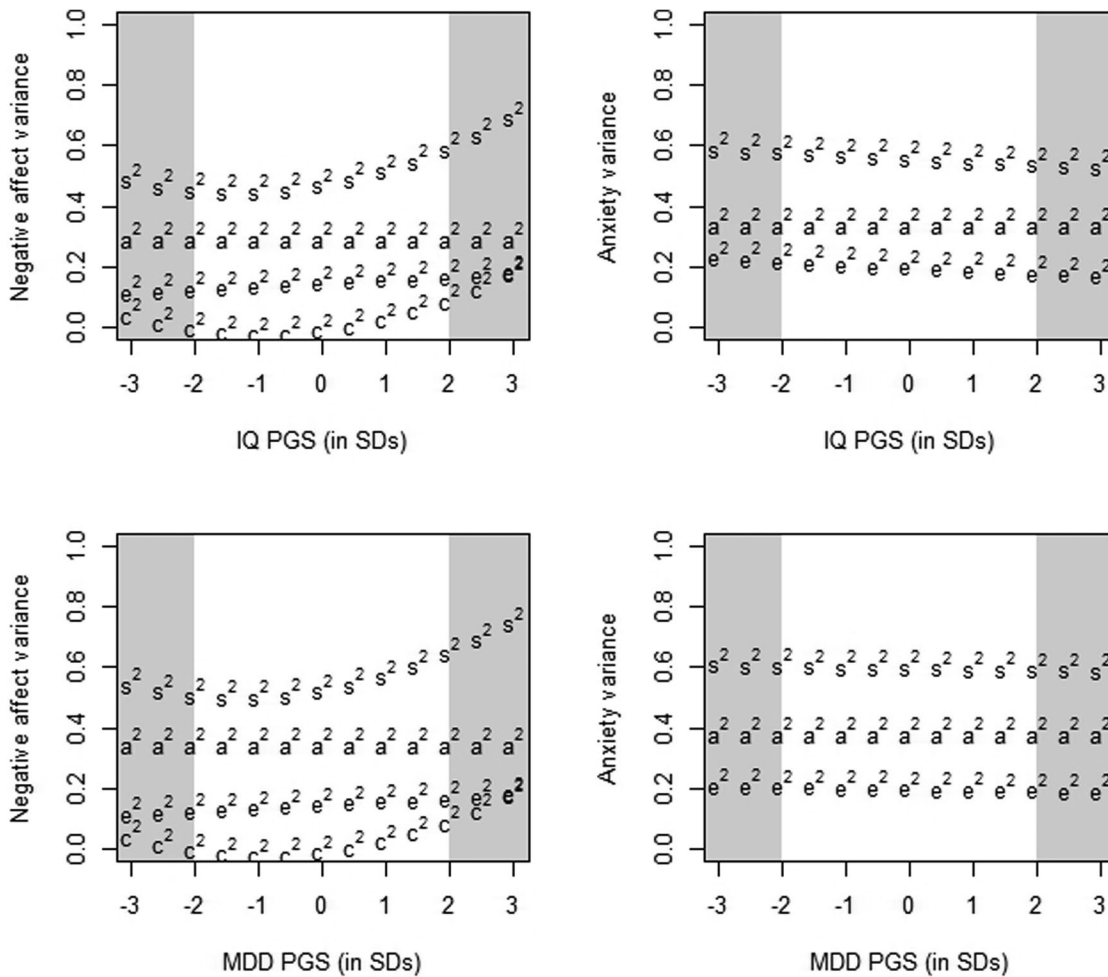


Figure 5. Genetic, environmental, and total variance of negative affect (left panes) and anxiety (right panes), conditional on the PGSs of IQ (top panes) and MDD (bottom panes). a^2 = total additive genetic variance ($a_L^2 + a_P^2$), c^2 = shared environmental variance, e^2 = unshared environmental variance, and s^2 = total phenotypic variance. The variance components are based on the parameter estimates from the analysis without the censoring correction. The shaded blocks indicate a PGS \pm 2-3 SD's from the mean.

advantage of the present approach to analyzing genotype-environment interaction based on measured PGSs is that it was relatively straightforward to incorporate in psychometric modeling.

A further possibility to improve the modeling of interaction effects on poorly-scaled measurements would be to use the standard errors of factor scores to moderate the residual variance of the trait. This

approach has shown advantages in the context of factor analysis and seems appropriate to address scaling issues in studies of $G \times E$ interaction. Simulations are underway to explore this conjecture.

Article Information

Conflict of Interest Disclosures: Each author signed a form for disclosure of potential conflicts of interest. MCN was funded by NIH grant R01DA049867.

Ethical Principles: The authors affirm having followed professional ethical guidelines in preparing this work. These guidelines include obtaining informed consent from human participants, maintaining ethical treatment and respect for the rights of human or animal participants, and ensuring the privacy of participants and their data, such as ensuring that individual participants cannot be identified in reported results.

Funding: This work was supported by the Royal Netherlands Academy of Science Professor Award under grant PAH/6635. SB is also supported by the Consortium on Individual Development (NWO 0240-001-003). The Netherlands Twin Register is supported by grant NWO 480-15-001/674: Netherlands Twin Registry Repository: researching the interplay between genome and environment, the Avera Institute for Human Genetics and by multiple grants from the Netherlands Organization for Scientific Research (NWO). MCN was supported by NIH grant R01DA049867.

Role of the Funders/Sponsors: None of the funders of this research had any role in the design and conduct of the study; collection, management, analysis, and interpretation of data; preparation, review, or approval of the manuscript; or decision to submit the manuscript for publication.

Acknowledgments: The ideas expressed in this manuscript are those of the authors alone, and endorsement by the authors' institutions or the funding agencies is not intended and should not be inferred. The authors would like to thank all twins and their family members for participating in the Netherlands Twin Register.

ORCID

Susanne Bruins  <http://orcid.org/0000-0001-6808-7075>
 Jouke-Jan Hottenga  <http://orcid.org/0000-0002-5668-2368>

Michael C. Neale  <http://orcid.org/0000-0002-7674-1767>
 René Pool  <http://orcid.org/0000-0001-5579-0933>
 Dorret I. Boomsma  <http://orcid.org/0000-0002-7099-7972>

Conor V. Dolan  <http://orcid.org/0000-0002-2496-8492>

Data availability statement

The Netherlands Twin Register data are stored in a repository and access to data may be provided for research purposes, after submitting a data request, which is evaluated by

a data access committee (DAC). Detailed information about working with NTR data and the NTR repository can be found at <https://ntr-data-request.psy.vu.nl/>

References

- Achenbach, T. M. (2001). *Manual for ASEBA school-age forms & profiles*. Research Center for Children, Youth & Families.
- Achenbach, T. M., Dumenci, L., & Rescorla, L. A. (2003). DSM-oriented and empirically based approaches to constructing scales from the same item pools. *Journal of Clinical Child and Adolescent Psychology: The Official Journal for the Society of Clinical Child and Adolescent Psychology, American Psychological Association, Division 53*, 32(3), 328–340. https://doi.org/10.1207/S15374424JCP3203_02
- Analytics, R., & Weston, S. (2014). doSNOW: Foreach parallel adaptor for the snow package. *R Package Version, 1*, 12.
- Balbona, J. V., Kim, Y., & Keller, M. C. (2021). Estimation of parental effects using polygenic scores. *Behavior Genetics, 51*(3), 264–278. <https://doi.org/10.1007/s10519-020-10032-w>
- Bates, T. C., Maher, B. S., Medland, S. E., McAloney, K., Wright, M. J., Hansell, N. K., Kendler, K. S., Martin, N. G., & Gillespie, N. A. (2018). The nature of nurture: Using a virtual-parent design to test parenting effects on children's educational attainment in genotyped families. *Twin Research and Human Genetics: The Official Journal of the International Society for Twin Studies, 21*(2), 73–83. <https://doi.org/10.1017/thg.2018.11>
- Boomsma, D. I., E de, G., van, B. G., & Koopmans, J. R. (1999). A religious upbringing reduces the influence of genetic factors on disinhibition: Evidence for interaction between genotype and environment on personality. *Twin Research, 2*(2), 115–125. <https://doi.org/10.1375/twin.2.2.115>
- Boomsma, D. I., & Martin, N. G. (2002). Gene-environment interactions. In H. D'Haenen, J. A. den Boer, & P. Wilner (Eds.), *Biological Psychiatry*, (Vol. 2, pp. 181–187). John Wiley & Sons.
- Boomsma, D. I., & Molenaar, P. C. M. (1986). Using LISREL to analyze genetic and environmental covariance structure. *Behavior Genetics, 16*(2), 237–250. <https://doi.org/10.1007/BF01070799>
- Anttila, V., Bulik-Sullivan, B., Finucane, H. K., Walters, R. K., Bras, J., Duncan, L., Escott-Price, V., Falcone, G. J., Gormley, P., Malik, R., Patsopoulos, N. A., Ripke, S., Wei, Z., Yu, D., Lee, P. H., Turley, P., Grenier-Boley, B., Chouraki, V., Kamatani, Y., ... Neale, B. M. (2018). Analysis of shared heritability in common disorders of the brain. *Science, 360*(6395), eaap8757. <https://doi.org/10.1126/science.aap8757>
- Bruins, S., Franić, S., Dolan, C. V., et al. (2022a). Structural equation modeling in genetics. *Handbook of Structural Equation Modeling*, Second Edition. Guilford Press.
- Bruins, S., van Bergen, E., Masselink, M. W., Barzeva, S. A., Hartman, C. A., Otten, R., Nanda N. J., Rommelse, N. N. J., Dolan, C. V., & Boomsma, D. I. (2022b). Are genetic and environmental risk factors for psychopathology

- amplified in children with below-average intelligence? A Population-Based Twin Study. *PsyArXiv*. <https://doi.org/10.31234/osf.io/tkg4u>
- Calaway, R., Weston, S., & Calaway, M. R. (2015). Package ‘foreach.’ R package 1–10.
- Choi, S. W., Mak, T. S.-H., & O’Reilly, P. F. (2020). Tutorial: A guide to performing polygenic risk score analyses. *Nature Protocols*, *15*(9), 2759–2772. <https://doi.org/10.1038/s41596-020-0353-1>
- de Zeeuw, E. L., Kan, K.-J., van Beijsterveldt, C. E. M., Mbarek, H., Hottenga, J.-J., Davies, G. E., Neale, M. C., Dolan, C. V., & Boomsma, D. I. (2019). The moderating role of SES on genetic differences in educational achievement in the Netherlands. *Npj Science of Learning*, *4*(1), 8. <https://doi.org/10.1038/s41539-019-0052-2>
- Dick, D. M., Viken, R., Purcell, S., Kaprio, J., Pulkkinen, L., & Rose, R. J. (2007). Parental monitoring moderates the importance of genetic and environmental influences on adolescent smoking. *Journal of Abnormal Psychology*, *116*(1), 213–218. <https://doi.org/10.1037/0021-843X.116.1.213>
- Dolan, C. V., Huijskens, R. C. A., Minică, C. C., Neale, M. C., & Boomsma, D. I. (2021). Incorporating polygenic risk scores in the ACE Twin Model to Estimate A–C Covariance. *Behavior Genetics*, *51*(3), 237–249. <https://doi.org/10.1007/s10519-020-10035-7>
- Eaves, L. J. (1977). Inferring the causes of human variation. *Journal of the Royal Statistical Society: Series A (General)*, *140*(3), 324–349. <https://doi.org/10.2307/2344924>
- Eaves, L. J. (2006). Genotype \times environment interaction in psychopathology: Fact or artifact? *Twin Research and Human Genetics : The Official Journal of the International Society for Twin Studies*, *9*(1), 1–8. <https://doi.org/10.1375/twin.9.1.1>
- Hagenbeek, F. A., Hirzinger, J. S., Breunig, S., Bruins, S., Dmitry V. Kuznetsov, D. V., Schut, K., Odintsova, V. V., & Dorret I. Boomsma, D. I. (2023). Maximizing the value of twin studies in health and behaviour. *Nature Human Behaviour*, *7*, 849–860. <https://doi.org/10.1038/s41562-023-01609-6>
- Hicks, B. M., South, S. C., Dirago, A. C., Iacono, W. G., & McGue, M. (2009). Environmental adversity and increasing genetic risk for externalizing disorders. *Archives of General Psychiatry*, *66*(6), 640–648. <https://doi.org/10.1001/archgenpsychiatry.2008.554>
- Johnson, W., McGue, M., & Iacono, W. G. (2009). School performance and genetic and environmental variance in antisocial behavior at the transition from adolescence to adulthood. *Developmental Psychology*, *45*(4), 973–987. <https://doi.org/10.1037/a0016225>
- Jöreskog, K. G. (1973). Analysis of Covariance Structures. In: Krishnaiah PR (ed) *Multivariate Analysis-III*. Academic Press. pp 263–285
- Kong, A., Thorleifsson, G., Frigge, M. L., Vilhjalmsdottir, B. J., Young, A. I., Thorgeirsson, T. E., Benonisdottir, S., Oddsson, A., Halldorsson, B. V., Masson, G., Gudbjartsson, D. F., Helgason, A., Bjornsdottir, G., Thorsteinsdottir, U., & Stefansson, K. (2018). The nature of nurture: Effects of parental genotypes. *Science*, *359*(6374), 424–428. <https://doi.org/10.1126/science.aan6877>
- Koopmans, J. R., Slutske, W. S., van Baal, G. C. M., & Boomsma, D. I. (1999). The Influence of Religion on Alcohol Use Initiation: Evidence for genotype \times environment interaction. *Behavior Genetics*, *29*(6), 445–453. <https://doi.org/10.1023/A:1021679005623>
- Middeldorp, C. M., Birley, A. J., Cath, D. C., Gillespie, N. A., Willemsen, G., Statham, D. J., de Geus, E. J. C., Andrews, J. G., van Dyck, R., Beem, A. L., Sullivan, P. F., Martin, N. G., & Boomsma, D. I. (2005). Familial clustering of major depression and anxiety disorders in Australian and Dutch twins and siblings. *Twin Research and Human Genetics*, *8*(6), 609–615. <https://doi.org/10.1375/twin.8.6.609>
- Molenaar, D., & Dolan, C. V. (2014). Testing systematic genotype by environment interactions using item level data. *Behavior Genetics*, *44*(3), 212–231. <https://doi.org/10.1007/s10519-014-9647-9>
- Molenaar, D., Middeldorp, C. M., Willemsen, G., Ligthart, L., Nivard, M. G., & Boomsma, D. I. (2016). Evidence for gender-dependent genotype by environment interaction in adult depression. *Behavior Genetics*, *46*, 59–71. <https://doi.org/10.1007/s10519-015-9752-4>
- Molenaar, D., van der Sluis, S., Boomsma, D. I., Haworth, C. M. A., Hewitt, J. K., Martin, N. G., Plomin, R., Wright, M. J., & Dolan, C. V. (2013). Genotype by environment interactions in cognitive ability: A survey of 14 studies from four countries covering four age groups. *Behavior Genetics*, *43*(3), 208–219. <https://doi.org/10.1007/s10519-012-9581-7>
- Molenaar, D., van der Sluis, S., Boomsma, D. I., & Dolan, C. V. (2012). Detecting specific genotype by environment interactions using marginal maximum likelihood estimation in the classical twin design. *Behavior Genetics*, *42*(3), 483–499. <https://doi.org/10.1007/s10519-011-9522-x>
- Molenaar, P., & Boomsma, D. I. (1987). Application of non-linear factor analysis to genotype-environment interaction. *Behavior Genetics*, *17*(1), 71–80. <https://doi.org/10.1007/BF01066011>
- Molenaar, P., Boomsma, D. I., & Dolan, C. V. (1999). The detection of genotype–environment interaction in longitudinal genetic models.
- Molenaar, P. C., Boomsma, D. I., Neeleman, D., & Dolan, C. V. (1990). Using factor scores to detect G \times E interactive origin of “pure” genetic or environmental factors obtained in genetic covariance structure analysis. *Genetic Epidemiology*, *7*(1), 93–100. <https://doi.org/10.1002/gepi.1370070116>
- Neale, M. C., Hunter, M. D., Pritikin, J. N., Zahery, M., Brick, T. R., Kirkpatrick, R. M., Estabrook, R., Bates, T. C., Maes, H. H., & Boker, S. M. (2016). OpenMx 2.0: Extended structural equation and statistical modeling. *Psychometrika*, *81*(2), 535–549. <https://doi.org/10.1007/s11336-014-9435-8>
- Nivard, M. G., Dolan, C. V., Kendler, K. S., Kan, K. J., Willemsen, G., van Beijsterveldt, C. E. M., Lindauer, R. J. L., van Beek, J. H. D. A., Geels, L. M., Bartels, M., Middeldorp, C. M., & Boomsma, D. I. (2015). Stability in symptoms of anxiety and depression as a function of genotype and environment: A longitudinal twin study from ages 3 to 63 years. *Psychological Medicine*, *45*(5), 1039–1049. <https://doi.org/10.1017/S003329171400213X>

- Purcell, S. (2002). Variance components models for gene-environment interaction in twin analysis. *Twin Research : The Official Journal of the International Society for Twin Studies*, 5(6), 554–571. <https://doi.org/10.1375/136905202762342026>
- R Core Team (2018). R: A language and environment for statistical computing
- Rathouz, P. J., Van Hulle, C. A., Rodgers, J. L., Waldman, I. D., & Lahey, B. B. (2008). Specification, testing, and interpretation of gene-by-measured-environment interaction models in the presence of gene-environment correlation. *Behavior Genetics*, 38(3), 301–315. <https://doi.org/10.1007/s10519-008-9193-4>
- Samejima, F. (1997). Graded response model. *Handbook of modern item response theory*. Springer. pp 85–100
- Samek, D. R., Hicks, B. M., Keyes, M. A., Bailey, J., McGue, M., & Iacono, W. G. (2015). Gene-environment interplay between parent-child relationship problems and externalizing disorders in adolescence and young adulthood. *Psychological Medicine*, 45(2), 333–344. <https://doi.org/10.1017/S0033291714001445>
- Samek, D. R., Hicks, B. M., Keyes, M. A., Iacono, W. G., & McGue, M. (2017). Antisocial peer affiliation and externalizing disorders: Evidence for gene \times environment \times development interaction. *Development and Psychopathology*, 29(1), 155–172. <https://doi.org/10.1017/S0954579416000109>
- Savage, J. E., Jansen, P. R., Stringer, S., Watanabe, K., Bryois, J., de Leeuw, C. A., Nagel, M., Awasthi, S., Barr, P. B., Coleman, J. R. I., Grasby, K. L., Hammerschlag, A. R., Kaminski, J. A., Karlsson, R., Krapohl, E., Lam, M., Nygaard, M., Reynolds, C. A., Trampush, J. W., ... Posthuma, D. (2018). Genome-wide association meta-analysis in 269,867 individuals identifies new genetic and functional links to intelligence. *Nature Genetics*, 50(7), 912–919. <https://doi.org/10.1038/s41588-018-0152-6>
- Selzam, S., Ritchie, S. J., Pingault, J.-B., Reynolds, C. A., O'Reilly, P. F., & Plomin, R. (2019). Comparing within- and between-family polygenic score prediction. *American Journal of Human Genetics*, 105(2), 351–363. <https://doi.org/10.1016/j.ajhg.2019.06.006>
- Tam, V., Patel, N., Turcotte, M., Bossé, Y., Paré, G., & Meyre, D. (2019). Benefits and limitations of genome-wide association studies. *Nature Reviews. Genetics*, 20(8), 467–484. <https://doi.org/10.1038/s41576-019-0127-1>
- van Beijsterveldt, C. E. M., Groen-Blokhuis, M., Hottenga, J. J., Franić, S., Hudziak, J. J., Lamb, D., Huppertz, C., de Zeeuw, E., Nivard, M., Schutte, N., Swagerman, S., Glasner, T., van Fulpen, M., Brouwer, C., Stroet, T., Nowotny, D., Ehli, E. A., Davies, G. E., Scheet, P., ... Boomsma, D. I. (2013). The Young Netherlands Twin Register (YNTR): Longitudinal twin and family studies in over 70,000 children. *Twin Research and Human Genetics : The Official Journal of the International Society for Twin Studies*, 16(1), 252–267. <https://doi.org/10.1017/thg.2012.118>
- van den Berg, S. M., Glas, C. A., & Boomsma, D. I. (2007). Variance decomposition using an IRT measurement model. *Behavior Genetics*, 37(4), 604–616. <https://doi.org/10.1007/s10519-007-9156-1>
- Van der Sluis, S., Posthuma, D., & Dolan, C. V. (2012). A note on false positives and power in GxE modelling of twin data. *Behavior Genetics*, 42(1), 170–186. <https://doi.org/10.1007/s10519-011-9480-3>
- Venables, W. N., & Ripley, B. D. (2013). *Modern applied statistics with S-PLUS*. Springer Science & Business Media.
- Verhulst, B., Prom-Wormley, E., Keller, M., Medland, S., & Neale, M. C. (2019). Type I error rates and parameter bias in multivariate behavioral genetic models. *Behavior Genetics*, 49(1), 99–111. <https://doi.org/10.1007/s10519-018-9942-y>
- Vilhjálmsdóttir, B. J., Yang, J., Finucane, H. K., Gusev, A., Lindström, S., Ripke, S., Genovese, G., Loh, P.-R., Bhatia, G., Do, R., Hayeck, T., Won, H.-H., Kathiresan, S., Pato, M., Pato, C., Tamimi, R., Stahl, E., Zaitlen, N., Pasaniuc, B., ... Price, A. L. (2015). Modeling linkage disequilibrium increases accuracy of polygenic risk scores. *American Journal of Human Genetics*, 97(4), 576–592. <https://doi.org/10.1016/j.ajhg.2015.09.001>
- Wagenmakers, E.-J., Krypotos, A.-M., Criss, A. H., & Iverson, G. (2012). On the interpretation of removable interactions: A survey of the field 33 years after Loftus. *Memory & Cognition*, 40(2), 145–160. <https://doi.org/10.3758/s13421-011-0158-0>
- Wesseldijk, L. W., Fedko, I. O., Bartels, M., Nivard, M. G., van Beijsterveldt, C. E. M., Boomsma, D. I., & Middeldorp, C. M. (2017). Psychopathology in 7-year-old children: Differences in maternal and paternal ratings and the genetic epidemiology. *American Journal of Medical Genetics. Part B, Neuropsychiatric Genetics : The Official Publication of the International Society of Psychiatric Genetics*, 174(3), 251–260. <https://doi.org/10.1002/ajmg.b.32500>
- Willoughby, E. A., Polderman, T. J. C., & Boutwell, B. B. (2023). Behavioural genetics methods. *Nature Reviews Methods Primers*, 3(1), 16. <https://doi.org/10.1038/s43586-022-00191-x>
- Winkler, T. W., Day, F. R., Croteau-Chonka, D. C., Wood, A. R., Locke, A. E., Mägi, R., Ferreira, T., Fall, T., Graff, M., Justice, A. E., Luan, J., Gustafsson, S., Randall, J. C., Vedantam, S., Workalemahu, T., Kilpeläinen, T. O., Scherag, A., Esko, T., Kutalik, Z., Heid, I. M., & Loos, R. J. F. (2014). Quality control and conduct of genome-wide association meta-analyses. *Nature Protocols*, 9(5), 1192–1212. <https://doi.org/10.1038/nprot.2014.071>
- Wray, N. R., Lin, T., Austin, J., McGrath, J. J., Hickie, I. B., Murray, G. K., & Visscher, P. M. (2021). From basic science to clinical application of polygenic risk scores: A primer. *JAMA Psychiatry*, 78(1), 101–109. <https://doi.org/10.1001/jamapsychiatry.2020.3049>
- Wray, N. R., Ripke, S., Mattheisen, M., Trzaskowski, M., Byrne, E. M., Abdellaoui, A., Adams, M. J., Agerbo, E., Air, T. M., Andlauer, T. M. F., Bacanu, S.-A., Bækvad-Hansen, M., Beekman, A. F. T., Bigdeli, T. B., Binder, E. B., Blackwood, D. R. H., Bryois, J., Buttenschøn, H. N., Bybjerg-Grauholm, J., ... Sullivan, P. F. (2018). Genome-wide association analyses identify 44 risk variants and refine the genetic architecture of major depression. *Nature Genetics*, 50(5), 668–681. <https://doi.org/10.1038/s41588-018-0090-3>