Genetically Informed Regression Analysis: Application to Aggression Prediction by Inattention and Hyperactivity in Children and Adults

Dorret I. Boomsma1,2 · Toos C. E. M. van Beijsterveldt1 · Veronika V. Odintsova1,2 · Michael C. Neale3 · Conor V. Dolan1,2

Received: 12 July 2020 / Accepted: 6 October 2020 / Published online: 1 December 2020 © The Author(s) 2020

Abstract
We present a procedure to simultaneously fit a genetic covariance structure model and a regression model to multivariate data from mono- and dizygotic twin pairs to test for the prediction of a dependent trait by multiple correlated predictors. We applied the model to aggressive behavior as an outcome trait and investigated the prediction of aggression from inattention (InA) and hyperactivity (HA) in two age groups. Predictions were examined in twins with an average age of 10 years (11,345 pairs), and in adult twins with an average age of 30 years (7433 pairs). All phenotypes were assessed by the same, but age-appropriate, instruments in children and adults. Because of the different genetic architecture of aggression, InA and HA, a model was fitted to these data that specified additive and non-additive genetic factors (A and D) plus common and unique environmental (C and E) influences. Given appropriate identifying constraints, this ADCE model is identified in trivariate data. We obtained different results for the prediction of aggression in children, where HA was the more important predictor, and in adults, where InA was the more important predictor. In children, about 36% of the total aggression variance was explained by the genetic and environmental components of HA and InA. Most of this was explained by the genetic components of HA and InA, i.e., 29.7%, with 22.6% due to the genetic component of HA. In adults, about 21% of the aggression variance was explained. Most was this was again explained by the genetic components of InA and HA (16.2%), with 8.6% due to the genetic component of InA.

Keywords Inattention · Hyperactivity · Aggression · Genetic and environmental prediction · Regression · Structural equation model

Introduction
Multivariate extensions of the classical twin design, that rest on trait and cross-trait comparisons of resemblances in mono- and dizygotic (MZ and DZ) twins, allow for inferences regarding pleiotropy and correlated environmental effects (Martin and Eaves 1977), the direction of causation between correlated traits (Heath et al. 1993; Duffy and Martin 1994), the moderation of genetic and environmental effects (Purcell 2002), and the analysis of the dimensionality of psychological (psychometric) instruments (Franić et al. 2013). In this contribution, we present an extension of the multivariate twin design that we developed to address questions about prediction of an outcome trait by multiple correlated variables. The model we present involves simultaneously fitting a multivariate genetic covariance structure model to estimate genetic (A, D) and environmental (C, E) variance matrices, and conducting the regression analyses.
Behavior Genetics (2021) 51:250–263

based on the genetic (A, D, or A + D) and environmental (C, E, or C + E) covariances.

We applied the model to measures of aggression, inattention, and hyperactivity that were collected in Dutch twins aged approximately 10 years, and in adult twins. Earlier research has indicated that aggression in children is influenced by genetic and common environmental factors (e.g., Porsch et al. 2016), while measures of attention deficit-hyperactivity disorder (ADHD), inattention, and hyperactivity tend to show strong evidence of non-additive genetic (dominance) influences (Derks et al. 2009). We therefore also investigated the conditions under which a variance decomposition model with both genetic dominance and common environmental influences could be fitted to the data. In univariate applications of the classical twin design, it is hardly ever possible to identify the contributions of both common environmental (C) factors and genetic dominance (D). Ozaki et al. (2011) presented an ACDE model using non-normal Structural Equation Modeling (nnSEM), that includes higher order moments as well as 1st- and 2nd-order moments, in which identification is achieved when not all four (ADCE) latent factors are distributed normally. We focus on identification in multivariate twin data, where identifying constraints can be formulated which allow for estimation of contributions from D and C factors in addition to A and E factors.

We considered the prediction of aggressive behavior by two dimensions of ADHD. ADHD is a neurobiological disorder that is characterized by symptoms of inattention and of hyperactivity/impulsivity, which may manifest in children and in adults. In children, positive associations have been found between broadly defined quantitative measures of aggression and ADHD and attention problems (Biederman et al. 1991; Jensen et al. 1997; Connor et al. 2010; Bartels et al. 2018, see: https://www.action-euproject.eu/ComorbidityChildAggression), and negative associations with academic performance (e.g., Hinshaw, 1992; Hinshaw et al. 2006; Vuoksimaa et al. 2020).

Individual differences in aggression and ADHD are strongly influenced by genetic factors (Derks et al. 2009; Hamshere et al. 2013; Faraone and Larsson, 2019; Odintsova et al. 2019). Studies of the etiology of the association between aggression and ADHD indicated that these associations were largely explained by pleiotropic genetic factors. Hur (2015) presented a review of twin studies on hyperactivity/inattention and Conduct Problems, which showed moderate to high (0.17–0.68) phenotypic correlations, and high genetic correlations (0.43–1.0). Based on a systematic review, Andersson et al. (2020) reported a genetic correlation between externalizing symptoms and ADHD of 0.49 (CI 0.37–0.61). These findings are consistent with the substantial genetic correlation between aggression and ADHD (rg = 1.00, SE = 0.07) that was estimated in a recent meta-analysis of genome-wide association studies of childhood aggression and ADHD (Ip et al. 2019).

Compared to the many studies on aggression and ADHD, a smaller number of analyses have focused on the relationship between aggression and hyperactivity/impulsivity or aggression and inattention. A comprehensive literature review and meta-analysis of studies in children, adolescents and adults on ADHD symptom dimensions indicated that aggressive behavior, and more generally externalizing disorders, are more strongly associated with hyperactivity/impulsivity than with inattention (Willcutt et al. 2012). The developmental trajectories of inattention and hyperactivity are different; young children are more likely to display hyperactive behaviors, while in middle childhood symptoms become more apparent and tend to persist into adulthood (Franke et al. 2018). Most studies of the association between aggression and ADHD subscales (see Willcutt et al. Supplementary Table 9) were done in children. The few publications in adults found no evidence that the associations of externalizing disorders with inattention and with hyperactivity differ.

Inattention and hyperactivity are not independent (e.g., Sokolova et al. 2016). Dolan et al. (2020) employed the classical twin design to analyze the correlation structure among measures of inattention and hyperactivity at the phenotype, genetic and environmental level. Inattention and hyperactivity were assessed by a variety of instruments. They concluded that the strong, broad-sense, genetic effects on inattention and hyperactivity are substantially correlated, regardless of instrument or rater.

Thus, when considering questions such as whether the association with aggression is stronger for inattention than for hyperactivity, we need to take into consideration that these two dimensions are not independent, e.g., there may be genetic pleiotropy, and that associations may differ across age groups. In this contribution, we investigated the differences between inattention and hyperactivity as predictors of aggression in a genetic design, analyzing data from MZ and DZ twins.

Method

Young Participants

The Young Netherlands Twin Register (YNTR) recruits newborn twins and multiples, and follows these children through development by survey studies and dedicated projects in subgroups (Boomsma et al. 2002; van Beijsterveldt et al. 2013; Ligthart et al. 2019). Recruitment of young twins began in 1987 and is ongoing. For the present study, we analyzed data on aggression, hyperactivity and inattention by maternal ratings of twins who were on average
10 years old (mean: 9.94 years; SD: 0.51). The twins were born between 1986 and 2006. In the YNTR, data on aggression were collected from 1995 onwards, and were available for all birth cohorts; data collection for hyperactivity and inattention began later, in 2001, so some twin pairs have incomplete phenotype information. There were 11,345 twin pairs (36% MZ). Table 2 summarizes the total number of participants and the number of missing data by twin member and phenotype.

**Adult Participants**

The Adult Netherlands Twin Register (ANTR) began longitudinal data collection by surveys in 1991 from adolescent and adult twins and their relatives. For the current study, we analyzed twin data from ANTR surveys 7 and 8, which were collected between 2004 and 2005 (survey 7), and between 2009 and 2011 (survey 8). The adult twins were on average 29.77 years old (SD: 12.5). The Conners’ Adult ADHD Rating Scales (CAARS), which we used to assess inattention and hyperactivity, was first introduced in ANTR survey 7 (Distel et al. 2007). However, this seventh survey did not include an assessment of aggression. ANTR survey 8 (Geels et al. 2013) was collected in two waves. Surveys from the first wave (83% of all responders) included the ASEBA-Adult Self Report aggression scale. The bottom part of Table 3 gives the total number of participants (total number of twin pairs is 7433; 46% MZ) and the number of missing values by twin member and phenotype. In contrast to the child data, the adult dataset had a substantial number of incomplete twin pairs (32% for MZ pairs and 51% for DZ pairs).

**Zygosity Assessment**

Most YNTR and ANTR surveys include a set of items concerning the twins’ physical resemblance and the degree to which the twins, in childhood, were confused by parents, other relatives, and strangers. In the YNTR and ANTR data, discriminant analyses were performed to assess the accuracy of zygosity classification based on survey items, using information from blood group and DNA polymorphisms as the index of true zygosity (Ligthart et al. 2019). In both the YNTR and ANTR, the accuracy of classification was high, ranging between 92 and 96%, depending on age and rater. In 31% of same-sex young twins and in the majority of same-sex ANTR twins (59%) zygosity assessment was based on DNA information.

**YNTR Phenotyping**

The Child Behavior Checklist (CBCL) is a standardized questionnaire designed for parents to report the frequency and intensity of their children’s behavioral and emotional problems (Achenbach et al. 2017). It is part of the Achenbach System of Empirically Based Assessment (ASEBA: https://aseba.org/), and consists of 120 items, which are rated on a 3-point scale. The response options range from ‘‘not true = 0’’, ‘‘somewhat or sometimes true = 1’’, to ‘‘very true or often true = 2’’. The Aggression Problem subscale contains 18 items; the total aggression score ranges from 0 to 36, allowing for up to 3 missing items (van Beijsterveldt et al. 2003). The Conners’ Parent Rating Scale-Revised (CPRS-R; Conners 2001; Conners et al. 1998; Derks et al. 2008) also assesses behavioral problems in children by parental ratings. The short version contains 27 items, which are rated on a 4-point scale, ranging from ‘‘not true at all = 0’’ to ‘‘very much true = 3’’. The two CPRS-R subscales that measure hyperactivity and inattention consist of 6 items each, allowing for 1 missing item per subscale. The phenotypic scores range from 0 to 18. The CPRS-R has good internal and test–retest reliability (Faries et al. 2001).

**ANTR Phenotyping**

The adult twins completed the ASEBA Adult Self Report (ASR) (Achenbach et al. 2017), which includes 15 aggression items that are rated on a 3-point scale, allowing for 3 missing items. The resulting aggression scores range from 0 to 30 (Hagenbeek et al. 2018). The Conners’ Adult ADHD Rating Scales screening self-report (CAARS—S:SV) includes two 9-item subscales for the quantitative assessment of inattentive symptoms (inattention) and hyperactive-impulsive symptoms (hyperactivity). There are no items common to the subscales. The items in the inattention and hyperactivity scales correspond to the symptoms that represent the diagnostic criteria of adult ADHD as outlined in DSM-IV-TR. All items were scored on a scale from ‘‘not true at all = 0’’ to ‘‘very much true = 3’’. The sum score of each subscale ranged from 0 to 27. Missing items were handled as per CAARS instructions (Conners et al. 1999; Saviouk et al. 2011) which allows the scoring of scales with up to two missing items.

**Statistical Analyses**

Statistical analyses were carried out in OpenMx (Neale et al. 2016) using full information maximum likelihood (ML) estimation. In all models, sex and age were included as fixed effects (the OpenMx scripts are given in the Online Appendix).
Phenotypic Regression of Aggression on Hyperactivity and Inattention

We first carried out the phenotypic analyses, in which we regressed aggression (Agg) on sex, age, hyperactivity (HA) and inattention (InA), in the child and in the adult cohort. The within-person regression model, depicted in Fig. 1 (where the fixed effects of sex and age are left out), is:

\[ \text{Agg}_i = b_0 + b_{\text{Sex}} \times \text{Sex}_i + b_{\text{Age}} \times \text{Age}_i + b_{\text{HA}} \times \text{HA}_i + b_{\text{InA}} \times \text{InA}_i + \varepsilon_i, \]

where the subscript \( i \) denotes person, and \( \varepsilon \) is the prediction error (regression residual). Conditional on sex and age, the phenotypic aggression variance \( s^2_{\text{Agg,sex}} \) was decomposed into four parts:

\[ s^2_{\text{Agg,sex}} = b^2_{\text{HA}} + s^2_{\text{HA}} + b^2_{\text{InA}} + s^2_{\text{InA}} + 2 \times b_{\text{HA}} \times b_{\text{InA}} \times s_{\text{HA,InA}} + s^2_{\varepsilon}. \]

The variance terms \( b^2_{\text{HA}} \) and \( s^2_{\text{InA}} \) can be attributed to hyperactivity and inattention, respectively. However, the term \( 2 \times b_{\text{HA}} \times b_{\text{InA}} \times s_{\text{HA,InA}} \) which arises when HA and InA are correlated, cannot be attributed unambiguously to either. We therefore report these three variance components separately. We fitted the phenotypic regression models in OpenMx (Neale et al. 2016) to the all twin data, regardless of the patterns of missingness. In these analyses we constrained the regression to be equal over MZ and DZ groups and over twin 1 and twin 2 (the two twins in a pair) and left the MZ twin 1–MZ twin 2 covariances and the DZ twin 1–DZ twin 2 covariances unconstrained, to accommodate the dependence of the MZ and DZ twin data (Neale et al. 1994). Below, we report the standardized regression coefficients, and the decomposition of the standardized variance of aggression.

Genetic Modeling: ADCE Twin Model

We first calculated the MZ and DZ \( 6 \times 6 \) phenotypic covariance matrices, whose standardized solution is given to describe MZ and DZ twin resemblances and fitted a trivariate Cholesky decomposition to the data (Brezinszki 2005) to estimate genetic and environmental covariance matrices for aggression, hyperactivity and inattention. The phenotypic \( (3 \times 3) \) covariance matrix of the phenotypes, conditional on age and sex \( (\Sigma_{\text{ph,age,sex}}) \) was decomposed into the following four covariance matrices (Martin and Eaves 1977; Franic et al. 2012):

In the 10 year olds:

\[ \Sigma_{\text{ph,age,sex}} = \Sigma_A + \Sigma_C + \Sigma_D + \Sigma_E, \]

and in the adults:

\[ \Sigma_{\text{ph,age,sex}} = \Sigma_A + \Sigma_D + \Sigma_E. \]

where \( \Sigma_A \) is the additive genetic covariance matrix, \( \Sigma_D \) the dominance genetic covariance matrix, \( \Sigma_C \) the common (shared between twins) environmental covariance matrix, and \( \Sigma_E \) the unique (unshared) environmental covariance matrix. To render the model identified in the 10-year olds, we added identifying constraints (informed by the MZ and DZ twin phenotypic correlations; see Tables 2 and 3). The \( 3 \times 3 \) covariance matrix \( \Sigma_C \), which was included to accommodate the contribution of shared environmental influences to aggression, was specified as follows:

\[ \Sigma_C = \Lambda_C \Lambda_C^t, \]

where \( t \) denotes transpose and \( \Lambda_C \)

\[
\begin{align*}
\Lambda_C & = \begin{pmatrix}
\text{Agg} & \text{HA} & \text{InA} \\
\text{Agg} & c_{11} & 0 & 0 \\
\text{HA} & c_{21} & 0 & 0 \\
\text{InA} & c_{31} & 0 & 0
\end{pmatrix}
\end{align*}
\]

The parameter \( c_{11} \) expresses the common environmental influences on aggression. We included the parameters \( c_{21} \) and \( c_{31} \) to accommodate shared environmental effects, if any, that are common to all three phenotypes. The \( 3 \times 3 \) covariance matrix \( \Sigma_D \) was modeled as \( \Lambda_D \Lambda_D^t \), where

\[
\begin{align*}
\Lambda_D & = \begin{pmatrix}
\text{Agg} & \text{HA} & \text{InA} \\
\text{Agg} & 0 & 0 & 0 \\
\text{HA} & d_{22} & 0 \\
\text{InA} & d_{32} & d_{33}
\end{pmatrix}
\end{align*}
\]
This is the Cholesky decomposition, with the dominance effects limited to hyperactivity and inattention. In both groups (children and adults), we modeled the additive genetic $3 \times 3$ covariance matrix $\Sigma_A$ and the unshared environmental covariance matrix $\Sigma_E$ as $\Lambda_A \Lambda_A^t$ and $\Lambda_E \Lambda_E^t$, respectively, where

$$
\Lambda_A = \begin{pmatrix}
    \text{Agg} & \text{HA} & \text{InA} \\
    a_{11} & 0 & 0 \\
    a_{21} & a_{22} & 0 \\
    a_{31} & a_{32} & a_{33}
\end{pmatrix}
$$

$$
\Lambda_E = \begin{pmatrix}
    \text{Agg} & \text{HA} & \text{InA} \\
    e_{11} & 0 & 0 \\
    e_{21} & e_{22} & 0 \\
    e_{31} & e_{32} & e_{33}
\end{pmatrix}
$$

That is, $\Lambda_A$ and $\Lambda_E$ were obtained from the full $3 \times 3$ Cholesky decomposition. The parameters were estimated by modeling the MZ and DZ twin covariance matrices ($6 \times 6$: 3 traits in two twins), conditional on age and sex:

$$
\Sigma_{\text{MZ twin 1}} = \Sigma_A + \Sigma_C + \Sigma_D + \Sigma_E \\
\Sigma_{\text{MZ twin 2}} = \Sigma_A + \Sigma_C + \Sigma_D + \Sigma_E \\
\Sigma_{\text{DZ twin 1}} = .5 \Sigma_A + \Sigma_C + .25 \Sigma_D \\
\Sigma_{\text{DZ twin 2}} = .5 \Sigma_A + \Sigma_C + .25 \Sigma_D
$$

where $\Sigma_A$, $\Sigma_C$, $\Sigma_D$ and $\Sigma_E$, are defined as above. We calculated the total genetic covariance matrix $\Sigma_G = \Sigma_A + \Sigma_D$ and the total environmental covariance matrix $\Sigma_T = \Sigma_C + \Sigma_E$ (in the adults, this is $\Sigma_T = \Sigma_E$).

### Genetic Modeling: A + D, C + E Regression Models

The trivariate genetic modeling provided an insight into the genetic and non-genetic correlations of hyperactivity and inattention with aggression, but did not address explicitly the question which of the subscales HA and InA is the stronger predictor of aggression. We included the regression of aggression on hyperactivity and inattention at the level of the genetic $\Sigma_G$ and the environmental covariance matrix $\Sigma_T$, where $\Sigma_G$ equals $\Sigma_A + \Sigma_D$, and $\Sigma_T$ equals $\Sigma_C + \Sigma_E$. In the adults, we have $\Sigma_T = \Sigma_E$. We did not attempt to conduct the regression analysis at the level of the individual (A, D, C and E) covariance matrices because $\Sigma_C$ and $\Sigma_D$ are positive semi-definite by definition (i.e., rank 1 and rank 2, respectively). In addition, $\Sigma_A$ was found to be positive semi-definite (rank 1) in the children. We therefore defined covariance matrices $\Sigma_G$ and $\Sigma_T$ to conduct the regression analyses at the total genetic (G) and the total (T) environmental level. Specifically, given $\Sigma_G$ (i.e., $\Sigma_A + \Sigma_D$),

$$
\Sigma_G = \begin{pmatrix}
    \text{Agg} & \text{HA} & \text{InA} \\
    \text{Agg}_2 & \text{Agg}_1 & \text{Agg}_0 \\
    \text{HA}_2 & \text{HA}_1 & \text{HA}_0 \\
    \text{InA}_2 & \text{InA}_1 & \text{InA}_0
\end{pmatrix}
$$

we partitioned the matrix into the following matrices:

$$
\Sigma_G = \begin{pmatrix}
    \text{Agg} & \text{HA} & \text{InA} \\
    \text{HA} & \text{InA} & \text{InA} \\
    \text{InA} & \text{InA} & \text{InA}
\end{pmatrix}
$$

i.e., the genetic covariance matrix of hyperactivity and inattention, and

$$
\Sigma_G = \begin{pmatrix}
    \text{Agg} & \text{HA} & \text{InA} \\
    \text{HA} & \text{InA} & \text{InA} \\
    \text{InA} & \text{InA} & \text{InA}
\end{pmatrix}
$$

i.e., the genetic covariance of aggression with hyperactivity and inattention. We calculated the genetic regression coefficients $b_G = [b_{G,HA}, b_{G,InA}]^t$ as follows: $b_G = \Sigma_G^{-1} \Sigma_G$. The decomposition of genetic variance associated with the genetic regression model is:

$$
n_{\text{G,Agg}}^2 = b_G^t \Sigma_G b_G + n_{\text{G,Agg}}^2 = b_{G,HA}^2 s_{G,HA}^2 + b_{G,InA}^2 s_{G,InA}^2 + 2 b_{G,HA} b_{G,InA} s_{G,InA,HA}^2 + s_{G,E}^2
$$
where $s_{\varepsilon}^2$ is the genetic prediction error variance. Using the same approach, we calculated $b_T = \left[ b_{T,HA}, b_{T,InA} \right]^T$ as follows $b_T = \Sigma_T^{-1} \Sigma_T$, and obtained the decomposition of total environmental variance:

$$s_{\varepsilon}^2 = b_T^T \Sigma_T^{-2} b_T + s_{\varepsilon, res} = b_{T,HA}^2 s_{HA}^2 + b_{T,InA}^2 s_{InA}^2 + 2* b_{T,HA} b_{T,InA} * s_{HA,InA} + s_{\varepsilon}^2.$$

Given estimates of the phenotypic, genetic, and environmental variance components, we standardized these by dividing by the total phenotypic, genetic, and environmental variance. The results of main interest are the genetic and environmental variance components standardized by the total phenotypic variance, as these reveal the relative contributions of genetic and environmental factors to the phenotypic regression of aggression on hyperactivity and inattention. Table 1 contains a summary of the decompositions of variance in the regression models.

### Results

#### Descriptive Statistics and Phenotypic Regression Analysis

Tables 2 and 3 contain the MZ and DZ correlation matrices, standard deviations, sample sizes and the number of missing values in the child and in the adult sample. In the children (Table 2), the phenotypic correlations among the three traits revealed that the correlation of Agg with HA (~0.60) is higher than the correlation of Agg with InA (~0.45), while the correlation between the two predictors InA and HA varies between ~0.29 and ~0.36 and the correlations of Agg with InA are consistently greater, between ~0.35 and ~0.48. Based on these results, it would seem that in the children HA may be the stronger predictor of Agg, while in the adults, InA is the stronger predictor. However, it is important to note that in adults hyperactivity and inattention were assessed four years before aggression was measured, which may have influenced the results.

We conducted phenotypic regression analyses on the basis of the within person phenotypic covariance (Agg-InA-HA) matrices. Results for these phenotypic analyses are summarized in Table 4A, which includes the standardized variance decomposition conditional on age and sex. Based on these results, Table 4B presents the proportions

| Table 1 | Variance components based on regression analysis with aggression as outcome and hyperactivity and inattention as predictors: variance components due to hyperactivity, inattention, the covariance of hyperactivity and inattention (unassigned to either hyperactivity or inattention) and a residual term |
|---------------------------------------|---------------------------------------|---------------------------------------|---------------------------------------|
| Hyperactivity (HA) | Inattention (InA) | Unassigned | Residual ($\varepsilon$) |
| $b_{T,HA}^2 s_{HA}^2 / s_{Agg}^2$ | $b_{T,InA}^2 s_{InA}^2 / s_{Agg}^2$ | $2* b_{T,HA} * b_{T,InA} * s_{HA,InA} / s_{Agg}^2$ | $s_{\varepsilon}^2 / s_{Agg}^2$ |
| Genetic variance components standardized by genetic variance $s_{G_{Agg}}^2$ |
| Hyperactivity (HA) | Inattention (InA) | Unassigned | Residual ($\varepsilon$) |
| $b_{G,HA}^2 s_{G_{HA}}^2 / s_{Agg}^2$ | $b_{G,InA}^2 s_{G_{InA}}^2 / s_{Agg}^2$ | $2* b_{G,HA} * b_{G,InA} * s_{G_{HA,InA}} / s_{Agg}^2$ | $s_{G_{\varepsilon}}^2 / s_{Agg}^2$ |
| Environmental variance components standardized by total environmental variance $s_{T_{Agg}}^2$ |
| Hyperactivity (HA) | Inattention (InA) | Unassigned | Residual ($\varepsilon$) |
| $b_{T,HA}^2 s_{T_{HA}}^2 / s_{Agg}^2$ | $b_{T,InA}^2 s_{T_{InA}}^2 / s_{Agg}^2$ | $2* b_{T,HA} * b_{T,InA} * s_{T_{HA,InA}} / s_{Agg}^2$ | $s_{T_{\varepsilon}}^2 / s_{Agg}^2$ |
| Genetic and environmental variance components standardized by phenotypic variance $s_{T_{Agg}}^2$ |
| Hyperactivity (HA) | Inattention (InA) | Unassigned | Residual ($\varepsilon$) |
| $b_{G,HA}^2 s_{G_{HA}}^2 / s_{Agg}^2$ | $b_{G,InA}^2 s_{G_{InA}}^2 / s_{Agg}^2$ | $2* b_{G,HA} * b_{G,InA} * s_{G_{HA,InA}} / s_{Agg}^2$ | $s_{G_{\varepsilon}}^2 / s_{Agg}^2$ |
| $b_{T,HA}^2 s_{T_{HA}}^2 / s_{Agg}^2$ | $b_{T,InA}^2 s_{T_{InA}}^2 / s_{Agg}^2$ | $2* b_{T,HA} * b_{T,InA} * s_{T_{HA,InA}} / s_{Agg}^2$ | $s_{T_{\varepsilon}}^2 / s_{Agg}^2$ |

$b = \text{regression coefficient, } s^2 = \text{total phenotypic variance, } s_G^2 = \text{(broad-sense) genetic variance, } s_T^2 = \text{total environmental variance}$
of explained phenotypic variance in aggression by the main effects of InA and HA, and by their covariance.

We first discuss the results in the children, where we note a consistent effect of sex. On average, girls scored lower than boys on all three phenotypes. The effects of age on Agg and InA were not significant (judging by the standard errors), but there was an effect of age on HA (α = 0.05). Even in this sample with limited variation in age (average age: 9.94 years; SD: 0.51), HA decreased with age, indicating fewer HA problems as children grow up (bAge = −0.235, se = 0.057).

Overall, sex and age combined explained about 1.9%, 3.6%, and 4.4% of the phenotypic variance of Agg, InA, and HA, respectively. Conditional on sex and age, we obtained regression coefficients of 0.165 (CI-95: 0.154–0.184 for InA) and 0.773 (CI-95: 0.758–0.798 for HA) in the regression of Agg on InA and HA. The total explained variance was 35.6%: InA explained 1.9% (CI-95: 1.6–2.4%) and HA explained 25.2% (CI-95: 23.6–26.7%) of the phenotypic Agg variance. The component due to the covariance between InA and HA explained an additional 8.5% (CI-95: 7.8–9.4%). Clearly HA emerged as the better phenotypic predictor, accounting for 25.2%/35.6% = ~71% of the explained variance, with InA accounting for 1.9%/35.6% = ~5% and the covariance of InA and HA accounting for 8.5%/35.6% = ~24% of the explained variance.

In the adults, we note a significant effect of sex on Agg and InA (α = 0.05), but no sex effect on HA. On average females scored higher on Agg and lower on InA. The effect of age was consistently negative, indicating lower scores with increasing age. Overall, sex and age combined explained 2.8%, 1.4%, and 0.1% of the phenotypic variance of Agg, InA, and HA, respectively. Conditional on sex and age, we obtained regression coefficients of 0.293 (CI-95: 0.273–0.300; InA) and 0.173 (CI-95: 0.138–0.174; HA) in the regression of Agg on InA and HA. The total explained variance was 18.6%: InA explained 10.6% (CI-95: 9.3–10.7%) and HA explained 3.1% (CI-95: 1.9–3.1%) of the phenotypic Agg variance. The component due to the covariance between InA and HA explained an additional 4.9% (CI-95: 4.4–5.1%). These results suggest that InA is the better phenotypic predictor in the adults, accounting to 10.6%/18.6% = 57% of the explained variance with HA and

Table 2  Children (10-year olds)

|        | Agg1 | InA1 | HA1 | Agg2 | InA2 | HA2 |
|--------|------|------|-----|------|------|-----|
| MZ     |      |      |     |      |      |     |
| Agg1   | 1.000|      |     |      |      |     |
| InA1   | 0.456| 1.000|     |      |      |     |
| HA1    | 0.605| 0.611| 1.000|      |      |     |
| Agg2   | 0.797| 0.381| 0.504| 1.000|      |     |
| InA2   | 0.385| 0.722| 0.483| 0.443| 1.000|     |
| HA2    | 0.518| 0.497| 0.772| 0.587| 0.620| 1.000|
| SD     | 4.772| 3.798| 2.905| 4.601| 3.829| 2.934|
| DZ     |      |      |     |      |      |     |
| Agg1   | 1.000|      |     |      |      |     |
| InA1   | 0.454| 1.000|     |      |      |     |
| HA1    | 0.587| 0.615| 1.000|      |      |     |
| Agg2   | 0.442| 0.230| 0.271| 1.000|      |     |
| InA2   | 0.232| 0.186| 0.227| 0.439| 1.000|     |
| HA2    | 0.254| 0.208| 0.276| 0.582| 0.620| 1.000|
| SD     | 4.865| 4.065| 3.081| 4.682| 4.072| 3.204|

Missing values (individuals) for twin1 and twin2

|        | Agg1 | InA1 | HA1 | Agg2 | InA2 | HA2 |
|--------|------|------|-----|------|------|-----|
| MZ     |      |      |     |      |      |     |
| Agg1   | 33   | 1125 | 1115|      | 1120 | 1113|
| InA1   | 33   | 1703 | 1696| 100  | 1727 | 1712|
| HA1    |      |      |     |      |      |     |
| Agg2   |      |      |     |      |      |     |
| InA2   |      |      |     |      |      |     |
| HA2    |      |      |     |      |      |     |
| SD     |      |      |     |      |      |     |

Total sample size (number of pairs) and marginal number of missing* values

|        | Agg1 | InA1 | HA1 | Agg2 | InA2 | HA2 |
|--------|------|------|-----|------|------|-----|
| MZ     |      |      |     |      |      |     |
| Agg1   | 4209 |      | 1125|      | 1115 | 1113|
| InA1   | 33   | 1703 | 1696| 100  |      |     |
| HA1    |      |      |     |      |      |     |
| Agg2   |      |      |     |      |      |     |
| InA2   |      |      |     |      |      |     |
| HA2    |      |      |     |      |      |     |
| SD     |      |      |     |      |      |     |

DZ and DZ correlation matrices of within-person, twin correlations (in bold) and cross-twin cross-trait correlations and standard deviations (corrected for sex and age) for Aggression (Agg), Inattention (InA) and hyperactivity (HA) in twin 1 and twin 2.

SD standard deviation

*In the earlier birth cohorts data on InA and HA were not collected.
the covariance of HA accounting for \(3.1\%/18.6\% = 17\%\) and \(4.9\%/18.6\% = 26\%\), respectively. However, the 4-year interval between the assessment of Agg and ADHD should be considered when interpreting these results.

**Combined Genetic Covariance Structure and Regression Analyses in 10-Year Olds**

Table 2 presents the correlations between twins and among scales in children. The MZ twin correlations were 0.79 (Agg), 0.72 (InA), and 0.77 (HA). The DZ correlations were substantially lower: 0.44 (Agg), 0.18 (InA), and 0.27 (HA). In the genetic covariance structure model, we therefore included additive genetic (A), dominance genetic (D), common environmental (C) and unique environmental (E) components (see model specification of matrices above). Based on twin data from mono- and dizygotic twins the full ADCE model is not identified and constraints as outlined above were applied to the \(3 \times 3\) C and D matrices. The estimates obtained in fitting the ADCE model are given in Table 5.

Heritability of Agg was 71% and common environment shared by twins accounted for 8% of the phenotype Agg variance. The estimates for the total heritability of HA and InA were also high, with relatively large contributions of genetic dominance. The broad-sense heritability of HA was 17% \((A) + 55\% (D) = 72\%\) and of InA 38% + 46% = 74%. Common environmental influences (shared with aggression) accounted for 1.4% and 1.1% of the variance of HA and InA, respectively. The C variance–covariance matrix is rank 1, which follows from our model specifications by definition, and matrix D is rank 2, again by definition. We found that matrix A is almost rank 1 (eigenvalues = 21.99, 0.006, ~ 0.00) and matrix E is rank 3. As mentioned above, given the ranks of these matrices, it is not possible to assess the regression model at the level of A, D and C (the covariance matrices of the predictors are singular). We therefore based the regression analyses on total genetic effects (broad sense: \(G = A + D\)) and the total environmental effects (\(T = C + E\)). The A + D and C + E matrices both have rank 3. The estimates of the regression coefficients, for
the phenotypic, genetic and non-genetic components in our model, are given in Table 6.

Effects of sex and age closely resembled the results as obtained in the phenotypic analysis (Table 4). The regression coefficients in the A + D part of the model were 0.134 for InA (CI-95: 0.091–0.174) and 0.830 for HA (CI = 95: 0.778–0.881). In the C + E part of the model these were 0.205 for InA (CI-95: 0.149–0.268) and 0.636 for HA

### Table 4 Phenotypic regression results in children and adults, with Aggression (Agg) as outcome and Inattention (InA) and Hyperactivity (HA) as predictors, where:

\[
\text{Agg}_i = b_0 + b_{\text{Sex}} \cdot \text{Sex}_i + b_{\text{Age}} \cdot \text{Age}_i + b_{\text{InA}} \cdot \text{InA}_i + b_{\text{HA}} \cdot \text{HA}_i + \epsilon_i
\]

A. Raw regression coefficients, with standard errors (se) or CI-95 intervals in parentheses

|                  | Agg   | InA   | HA    |
|------------------|-------|-------|-------|
| **Children**     |       |       |       |
| Intercept (b_0)  | 0.044 (se 0.890) | 5.23 (se 0.700) | 5.38 (se 0.567) |
| Sex (b_{Sex})    | -1.18 (se 0.063) | -1.47 (se 0.063) | -1.33 (se 0.047) |
| Age (b_{Age})    | 0.050 (se 0.089) | -0.086 (se 0.070) | -0.235 (se 0.057) |
| Phenotypic Agg on InA (b_{InA}) | 0.165 (0.154–0.184) |
| Phenotypic Agg on HA (b_{HA})  | 0.773 (0.758–0.798) |
| **Adults**       |       |       |       |
| Intercept (b_0)  | 3.86 (se 0.126) | 7.57 (se 1.34) | 7.58 (se 0.123) |
| Sex (b_{Sex})    | 0.727 (se 0.089) | -0.244 (se.093) | 0.050 (se 0.086) |
| Age (b_{Age})    | -0.036 (se 0.003) | -0.035 (se .003) | -0.010 (se 0.003) |
| Phenotypic Agg on InA (b_{InA}) | 0.293 (0.273–0.300) |
| Phenotypic Agg on HA (b_{HA})  | 0.173 (0.138–0.174) |

B. Proportions of explained phenotypic variance of Agg with CI-95 intervals in parentheses

|                  | \(b_{InA}^2 s^2_{InA}/s^2_{\text{Agg}}\) | \(b_{HA}^2 s^2_{HA}/s^2_{\text{Agg}}\) | \(2b_{HA}b_{InA} s_{HA,InA}/s^2_{\text{Agg}}\) |
|------------------|------------------------------------------|------------------------------------------|------------------------------------------|
| **Kids**         |                                          |                                          |                                          |
| Proportion of variance | 0.019 (0.016–0.024) | 0.252 (0.236–0.267) | 0.085 (0.078–0.094) |
| **Adults**       |                                          |                                          |                                          |
| Proportion of variance | 0.106 (0.093–0.107) | 0.031 (0.019–0.031) | 0.049 (0.044–0.051) |

\(s^2\) stands for variance and s represents standard deviation; \(s_{HA,InA}\) is the covariance of InA and HA

### Table 5 Estimates of additive and dominance genetic and common and unique environmental variance–covariance matrices (\(\Sigma\)), based on fitting an ADCE model to the trivariate twin data, conditional on age and sex

|                  | Agg   | InA   | HA   | Agg   | InA   | HA   | Agg   | InA   | HA   |
|------------------|-------|-------|------|-------|-------|------|-------|-------|------|
| **Child**        |       |       |      |       |       |      |       |       |      |
| \(\Sigma A\)     | 16.20 | 1     | 1    | –     | –     | –    | 1.86  | 1     | 1    |
| \(\Sigma D\)     | 6.66  | 2.77  | 1    | –     | 8.84  | 0.526| 0.66  | 0.23  | 1    |
| \(\Sigma C\)     | 7.00  | 2.89  | 3.02 | –     | 3.28  | 4.40 | 0.44  | 0.16  | 0.10 |
| \(\Sigma E\)     |       |       |      |       |       |      | 4.50  | 0.258 | 0.392|

|                  | Agg   | InA   | HA   | Agg   | InA   | HA   | Agg   | InA   | HA   |
|------------------|-------|-------|------|-------|-------|------|-------|-------|------|
| **Adult**        |       |       |      |       |       |      |       |       |      |
| \(\Sigma A\)     | 2.36  | 0.584 | 0.814| 2.94  | 0.546 | 0.814| 5.95  | 0.281 | 0.224|
| \(\Sigma D\)     | 1.88  | 4.36  | 0.831| 1.24  | 1.76  | -0.041| 1.91  | 7.69  | 0.402|
| \(\Sigma E\)     | 0.08  | 2.52  | 2.52 | 2.10  | -0.08 | 2.26 | 1.50  | 3.05  | 7.49 |

|                  | Agg   | InA   | HA   | Agg   | InA   | HA   |
|------------------|-------|-------|------|-------|-------|------|
| \(h^2\)         | 0.717 | 0.172 | 0.382|      | 0.550 | 0.462| 0.082 | 0.014 | 0.011|
| \(d^2\)         | 0.550 | 0.462 |      |      |      |      | 0.199 | 0.262 | 0.208|
| \(c^2\)         | 0.117 | 0.260 | 0.127| 0.190 | 0.528 | 0.556| 0.631 |

|                  | Agg   | InA   | HA   | Agg   | InA   | HA   |
|------------------|-------|-------|------|-------|-------|------|
| \(h^2\)         | 0.209 | 0.315 | 0.117| 0.260 | 0.127 | 0.190|
| \(d^2\)         | 0.315 | 0.260 |      |      |      |      | 0.528 | 0.556 | 0.631|

The lower triangles give (co)variances, and the upper triangle correlations. The last row (for children and for adults) contains the standardized variance components (\(h^2\) is heritability; \(d^2\), \(c^2\) and \(e^2\) give the proportions of variance explained by D, C and E)
Table 6: Genetic (G(A + D)) and environmental (T(C + E)) regression results in children, where for both the genetic and environmental regression $\text{Agg}_i = b_0 + b_{\text{Sex}} \cdot \text{Sex}_i + b_{\text{Age}} \cdot \text{Age}_i + b_{\text{InA}} \cdot \text{InA}_i + b_{\text{HA}} \cdot \text{HA}_i + \epsilon_i$

| Child | Agg | InA | HA |
|-------|-----|-----|-----|
| Intercept ($b_0$) | 0.047 (se 0.760) | 5.28 (se 0.705) | 5.39 (se 0.515) |
| Sex ($b_{\text{Sex}}$) | -1.18 (se 0.063) | -1.49 (se 0.063) | -1.33 (se 0.047) |
| Age ($b_{\text{Age}}$) | 0.050 (se 0.076) | -0.090 (se 0.071) | -0.237 (se 0.052) |
| G(A + D) on InA ($b_{\text{InA}}$) | 0.134 (0.091–0.174) | | |
| G(A + D) on HA ($b_{\text{HA}}$) | 0.830 (0.778–0.881) | | |
| T(C + E) on InA ($b_{\text{InA}}$) | 0.205 (0.149–0.268) | | |
| T(C + E) on HA ($b_{\text{HA}}$) | 0.636 (0.532–0.744) | | |

Proportions of explained phenotypic variance of Agg with CI-95 intervals in parentheses

- Explained by G(A + D): 0.414
- Explained by T (C + E): 0.219
- Explained by G (A + D): 0.297
- Explained by T (C + E): 0.062

The remaining 37% is distributed over the other 5 remaining sources of variance.

Combined Genetic Covariance Structure and Regression Analyses in Adults

Table 3 includes the correlations between twins and among scales in the adult twins. The MZ twin correlations were 0.46 (Agg), 0.42 (InA), and 0.36 (HA). The DZ correlations were substantially lower: 0.17 (Agg), 0.20 (InA), and 0.13 (HA). In the genetic covariance structure model, we therefore included additive genetic (A), dominance genetic (D), and unique environmental (E) components. The estimates obtained in fitting the ADE model are given in Table 5 (bottom). The narrow sense heritabilities are 21% (Agg), 31% (InA), and 12% (HA). The dominance variance components are relatively large: 26% (Agg), 13% (InA), and 19% (HA), giving rise to broad-sense heritabilities of 21 + 26 = 47% (Agg), 31 + 13 = 44% (InA), and 12 + 19 = 31% (HA). The unshared environmental variance is relatively large: 53% (Agg), 56% (InA), and 62% (HA). The estimates of the regression coefficients, for the phenotypic, genetic and non-genetic components in our model, are given in Table 7.

The effects of sex and age closely resembled the results of the phenotypic analysis (Table 4). The regression coefficients in the A + D part of the model were 0.399 for InA (CI-95: 0.374–0.485) and 0.276 for HA (CI = 95: 0.149–0.395). In the E part of the model these were 0.201...
for InA (CI-95: 0.169–0.256) and 0.118 for HA (CI-95: 0.055–0.176). Table 7 also contains the decomposition of standardized phenotype Agg variance based on the regression models. In the A + D part of the model, the total explained (broad-sense) genetic variance of Agg was 34.8%, which is decomposed into 18.4% (InA; CI-95: 10.8–26.5%), 6.2% (HA; CI-95: 5–11%), and 10.1% (CI-95: 6.8–12.9%), due to the broad-sense genetic covariance between InA and HA. In the E part of the model, the total explained environmental variance was 16.4%. This is decomposed into 5.2% (InA; CI-95: 5.2–6.9%), 1.7% (HA; CI-95: 1.6–4%), and 2.4% (CI-95: 1.3–2.9%) due to the environmental covariance of InA and HA. At the broad sense genetic and environmental levels, InA emerges as the better predictor (genetic: 18.4%/34.8% = 53%; environmental: 5.2%/16.4% = 32%).

To evaluate the predictive contributions to the phenotype variance of Agg, we standardized by the phenotypic variance (Table 7 bottom part). The total explained variance was 16.2% (A + D) + 4.9% (E) = 21%. As expected this resembles the percentage of the explained variance in the phenotypic analyses (as mentioned above: 18.6%). The 21% is decomposed as follows. A + D contributes 8.6% due to genetic InA, 2.9% due to genetic HA, 4.7% due to the genetic covariance of InA and HA. E contributes 2.7% due to environmental InA, 0.9% due to environmental HA, and 1.3% due to the environmental covariance between InA and HA. By far the best predictor is genetic InA, which accounts for 8.6%/21% = 41% of the phenotypic variance of Agg.

### Discussion

In this contribution, we integrated a regression model within genetic covariance modeling. We applied the model to data from children and adults to address the question of differential prediction of aggression (Agg) by two components of ADHD, i.e., inattention (InA) and hyperactivity (HA). These types of questions of a best genetic predictor of an outcome trait or disease may come up in multiple contexts, such as the prediction of educational attainment by cognitive ability and non-cognitive skills (Demange et al. 2020) or hypertension and cardiovascular outcomes by multiple correlated factors (Lucaroni et al. 2019).

Also, the integrated model that we presented can be applied beyond the classical twin design to any genetically informative dataset or design that allows estimation of genetic and a non-genetic covariance matrices, including adoption or family studies and single-nucleotide polymorphism (SNP) based approaches to infer heritability and genetic covariance matrices from GWA studies or their summary statistics. The possibility to consider ADCE models, rather than limiting to e.g., AE or ACE, depends on the study design and the appropriate identifying constraints. In our model for twin data these constraints involved specifying a one-factor model for the common environmental influences and the absence of genetic dominance for 1 of the 3 phenotype outcomes. We note that the present approach of estimating genetic and environmental covariance matrices, and simultaneously modeling these, differs slightly from

---

Table 7 Genetic (G(A + D)) and environmental (T(E)) regression results in adults, where for both the genetic and environmental regression $Agg_i = b_0 + b_{Sex} \times Sex_i + b_{Age} \times Age_i + b_{InA} \times InA_i + b_{HA} \times HA_i + \epsilon_i$

| Adults | Agg | InA | HA |
|--------|-----|-----|-----|
| Raw regression coefficients with standard errors (se) or CI-95 intervals in parentheses | | | |
| Intercept ($b_0$) | 3.86 (se 0.126) | | |
| Sex ($b_{Sex}$) | 0.722 (se 0.090) | | |
| Age ($b_{Age}$) | -0.003 (se 0.003) | | |
| G (A + D) Agg on InA ($b_{InA}$) | 0.399 (0.374–0.485) | | |
| G (A + D) Agg on HA ($b_{HA}$) | 0.276 (0.149–0.395) | | |
| T (E) Agg on InA ($b_{InA}$) | 0.201 (0.169–0.256) | | |
| T (E) Agg on HA ($b_{HA}$) | 0.118 (0.055–0.176) | | |

Proportions of explained phenotypic variance of Agg with CI-95 intervals in parentheses

| Explained by G(A + D) | Total genetic | | |
|-----------------------|---------------|---------------|
| $b_{G,InA}^2$ InA/s^2 Agg | $b_{G,HA}^2$ HA/s^2 Agg | $b_{G,InA}^2$ InA/s^2 Agg |
| Explained by T(E) | Total environ | | |
| $b_{T,InA}^2$ InA/s^2 Agg | $b_{T,HA}^2$ HA/s^2 Agg | $b_{T,InA}^2$ InA/s^2 Agg |
| Explained by G(A + D) | Total phenot | | |
| $b_{G,InA}^2$ InA/s^2 Agg | $b_{G,HA}^2$ HA/s^2 Agg | $b_{G,InA}^2$ InA/s^2 Agg |
| Explained by T(E) | | | |
| $b_{T,InA}^2$ InA/s^2 Agg | $b_{T,HA}^2$ HA/s^2 Agg | $b_{T,InA}^2$ InA/s^2 Agg |
| Explained by T(E) | Total genetic | Total environ | Total phenot |
| 0.497 | 0.027 (0.020–0.027) | 0.009 (0.008–0.010) | 0.013 (0.010–0.014) |
the standard multivariate genetic covariance structure mod-
eling where genetic and environmental covariance matrices 
are subjected directly to a structural equation model (e.g., 
a growth curve model, autoregressive model, or a common 
factor model). However, the present approach allows us to 
fit the model of interest (i.e., the regression model) to the 
broad-sense genetic (A + D) and the total environmental 
(C + E) covariance matrices, provided that these are posi-
tive definite.

Application of these methods produced a clear set of 
results concerning the prediction of aggression in children 
and in adults. In children, genetic hyperactivity was without 
doubt the stronger predictor of aggression, after taking into 
count the effects of inattention and the shared covariance 
of hyperactivity and inattention. A stronger predictive value 
of HA for aggression in children is consistent with several 
lines of research. There is evidence of different neural cor-
relates of ADHD with predominantly hyperactive-impulsive, 
previously inattentive and the combined subtype (Saad 
et al. 2017). Hyperactivity is a stronger predictor of con-
duct problems than inattention in girls with ADHD (Lee 
and Hinshaw 2006). The information on 10-year old twins 
was collected from maternal ratings, and rater effects could 
have contributed to the results that were obtained. Vierikko 
et al. (2004) assessed the relation between aggression and 
hyperactivity/impulsivity from parents in the home situa-
tion and from teachers in the classroom. They report high 
genetic correlations between aggression and hyperactivity 
both when analyses were based on teacher and on parental 
ratings.

The results obtained in the analyses of adult self-ratings 
led to different conclusions concerning the prediction of 
aggression: genetic inattention clearly was the better pre-
dictor of aggression, again after considering the effects of 
hyperactivity and the covariance of hyperactivity and inat-
tention. We note, however, that the adult dataset included 
self-report measures collected at different points in time. 
Still, assuming the test–retest reliability of the inattention 
test and the hyperactivity are about the same in adults, the 
difference in measurement occasion would not explain the 
relatively stronger role of inattention.

In conclusion, our genetic modeling of the trivariate twin 
data provided an insight into the genetic and non-genetic 
predictors of aggression. Standard Cholesky decompositions 
are commonly used to obtain estimates of genetic and envi-
ronmental covariance matrices. We note that the Cholesky 
parameterization itself can be used as a regression model 
(with the dependent variable as the last variable; e.g., de 
Jong 1999). However, the present approach has the advan-
tage of basing the regression model on the A + D and the 
C + E covariance matrices, which is useful if the covariances 
matrices (A and/or D, C and/or E) are (near) singular. In 
that case, we consider the option to be able to address the 
prediction issue at the broad-sense genetic or total environ-
mental level to be a worthwhile one. In addition, our pre-
sent approach to regression modeling in OpenMx allows 
a decomposition of the variance of the dependent variable 
(Aggression) into raw and standardized variance compo-
nents. We carried out the ACDE decomposition and regres-
sion analysis separately in children and adults, with age and 
sex as fixed covariates. We note that the present implement-
tion of the regression model in genetic covariance struc-
ture modeling can be extended to include fixed covariates as 
moderators of the regression parameters. For instance, in our 
adult data the mean age is ~30 years, but the variation in age 
is quite large (SD: 12.5). The present approach to the regres-
sion analysis can be extended to include age as a continuous 
moderator. We included the regression of aggression on HA 
and InA at the genetic and environmental level and obtained 
estimates of the phenotypic, genetic, and environmental vari-
ance components standardized by the total phenotypic vari-
ance, which revealed the relative contributions of genetic 
and environmental factors to the phenotypic regression of 
aggression on hyperactivity and inattention.

Acknowledgements We warmly thank all participants who donate 
their time to the Netherlands Twin Register research projects. We 
acknowledge funding from the Borderline Personality Disorder 
Research Foundation; ZonMW (31600008), the European Research 
Council (ERC–230374), NWO (480-15-001/674); Netherlands Twin 
Registry Repository; the European Union Seventh Framework Program 
(FP7/2007-2013) grant agreement 602768: “Aggression in Children: 
Unraveling gene-environment interplay to inform Treatment and Inter-
vention strategies” (ACTION); National Institute on Drug Abuse grant 
DA049867; and the KNAW Academy Professor Award (PAH/6635) 
to DIB. We thank Sofieke Kevenaar for her comments on the paper.

Compliance with Ethical Standards

Conflict of interest Dorret I. Boomsma, Toos C. E. M. van Beijs-
terveldt, Veronika Odintsova, Michael C. Neale, Conor V. Dolan de-
clare that they have no conflict of interest.

Ethical Approval All procedures performed in studies involving 
human participants were in accordance with the ethical standards of 
the institutional and/or national research committee and with the 1964 
Helsinki declaration and its later amendments or comparable ethical 
standards. The studies were approved by the Central Ethics Committee 
On Research Involving Human Subjects of the VU University Medical 
Center, Amsterdam, an Institutional Review Board certified by the U.S. 
Office of Human Research Protections (IRB number IRB-2991 under 
Federal-wide Assurance-3703).

Informed Consent Informed consent was obtained for all participants.

Open Access This article is licensed under a Creative Commons Attri-
bution 4.0 International License, which permits use, sharing, adapta-
tion, distribution and reproduction in any medium or format, as long 
as you give appropriate credit to the original author(s) and the source, 
provide a link to the Creative Commons licence, and indicate if changes 
were made. The images or other third party material in this article are 
included in the article’s Creative Commons licence, unless indicated
otherwise in a credit line to the material. If material is not included in
the article’s Creative Commons licence and your intended use is not
permitted by statutory regulation or exceeds the permitted use, you will
need to obtain permission directly from the copyright holder. To view a
copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

Achenbach TM, Ivanova MY, Rescorla LA (2017) Empirically based
assessment and taxonomy of psychopathology for ages 1½–90+
years: developmental, multi-informant, and multiculturalfindings.
Compr Psychiatry 79:4–18

Andersson A, Tuvblad C, Chen Q, Du Rietz E, Cortese S, Kujahalkola
R, Larsson HJ (2020) Research review: the strength of the genetic
overlap between ADHD and other psychiatric symptoms—a sys-
tematic review and meta-analysis. J Child Psychol Psychiatry.
https://doi.org/10.1111/jcpp.13233

Bartels M, Hendriks A, Mauri M, Krapohl E, Whipp A, Bolhuis K,
Conde LC, Luningham J, Ip HF, Hagenbeek F, Roetman P, Gatej
R, Lamers A, Nivard M, van Dongen J, Lu Y, Middeldorp C, van
Beijsterveldt T, Vermeiren R, Hankemeier T, Kluft C, Medland
S, Lundström S, Kung R, Pulkkinen L, Vuoksimaa E, Korhonen
T, Martin NG, Lurbe G, Finkenauer C, Fanos V, Tiemeier H,
Lichtenstein P, Plomin R, Kaprio J, Boomsma DI (2018) Child-
hood aggression and the co-occurrence of behavioural and emo-
tional problems: results across ages 3–16 years from multiple
raters in six cohorts in the EU-ACTION project. Eur Child Ado-
escence Psychiatry 27(9):1105–1112

Biederman J, Newcorn J, Sprich S (1991) Comorbidity of attention
deficit hyperactivity disorder with conduct, depressive, anxiety,
and other disorders. Am J Psychiatry 148:564–577

Boomsma DI, Vink JM, van Beijsterveldt TC, de Geus EJ, Borsboom
D, Br naik JJ, van Beijsterveldt CE, Mulder EJ, Derks EM, Riese H, Willemsen G, Bartels M, van
den Berg M, Kupper NH, Polderman TJ, Posthuma D, Rietveld
MJ, Stubble JH, Knol LI, Stroet T, van Baal GC (2002) Nether-
lands twin register: a focus on longitudinal research. Twin Res
5(5):401–406

Breznitz C (2005) La méthode de Cholesky. Revue d’Histoire des
Mathématiques 11(2):205–238

Conners CK, Erhardt D, Sparrow E (1999) Conners’ adult ADHD
rating scales (CAARS). Multi-Health systems Inc, North Ton-
awanda, NY, p 144

Conners CK (2001) Conners’ rating scales-revised. Multi-Health Sys-
tems Inc, NY, Toronto

Conners CK, Sitarenios G, Parker JD, Epstein JN (1998) The revised
conners’ parent rating scale (CPRS-R): factor structure, reliability,
and criterion validity. J Abnorm Child Psychol 26:257–268

Connor DF, Chartier KG, Preen EC, Kaplan RF (2010) Impulsive
aggression in attention-deficit/hyperactivity disorder: symptom
severity, co-morbidity, and attention-deficit/hyperactivity disorder
subtype. J Child Adolesc Psychopharmacol 20(2):119–126

de Jong PF (1999) Hierarchical regression analysis in structural equa-
tion modeling. Struct Eqn Model 6(2):198–211

Demange PA et al (2020) Parental influences on offspring education:
direct genetic effects of non-cognitive skills. bioRxiv. https://doi.
org/10.1101/2020.09.15.296236

Derks EM, Hudziak JJ, Dolan CV, van Beijsterveldt TC, Verhulst FC,
Boomsma DI (2008) Genetic and environmental influences on the
relation between attention problems and attention deficit hyperac-
tivity disorder. Behav Genet 38(1):11–23

Derks EM, Hudzik J, Boomsma DI (2009) Genetics of ADHD,
hyperactivity, and attention problems. In: Yong-Kyu K (ed) Hand-
book of behavior genetics. Springer, NY, pp 361–378

Distel MA, Ligthart L, Willemsen G, Nyholt DR, Truill TJ, Boomsma
DI (2007) Personality, health and lifestyle in a questionnaire fam-
ily study: a comparison between highly cooperative and less coop-
erative families. Twin Res Hum Genet 10(2):348–353

Dolan CV, de Zeeuw EL, Zayats T, van Beijsterveldt CEM, Boomsma
DI (2020) The (Broad-Sense) genetic correlations among four
measures of inattention and hyperactivity in 12 year olds. Behav
Genet 50(4):273–288

Duffy DL, Martin NG (1994) Inferring the direction of causation in
cross-sectional twin data: theoretical and empirical considera-
tions. Genet Epidemiol 11(6):483–502

Faraone SV, Larsson H (2019) Genetics of attention deficit hyperactiv-
ty disorder. Mol Psychiatry 24(4):562–575

Faries DE, Yalcin I, Harder D, Heiligenstein IH (2001) Validation of
the ADHD rating scale as a clinician administered and scored
instrument. J Atten Disord 5(2):107–115

Franic S, Dolan CV, Borsboom D, Boomsma DI (2012) Structural
equation modeling in genetics. In: Hoyle R (ed) Handbook of
structural equation modeling. Guilford Press, NY, pp 617–635

Franic S, Dolan CV, Borsboom D, Hudziak JJ, van Beijsterveldt CE,
Boomsma DI (2013) Can genetics help psychometrics? Improv-
ing dimensionality assessment through genetic factor modeling.
Psychol Methods 18(3):406–433

Franke B, Michelini G, Asherson P, Banaschewski T, Bilbow A, Buite-
laar JK, Cormand B, Faraone SV, Ginsberg Y, Haavik J, Kuntsi J,
Larsson H, Lesch KP, Ramos-Quiroga JA, Réthelyi JM, Riba-
Mees R, Rief A (2018) Live fast, die young? a review on the develop-
mental trajectories of ADHD across the lifespan. Eur Neuropsy-
chopharmacol 28(10):1059–1088

Geels LM, Vink JM, van Beek JH, Bartels M, Willemsen G, Boomsma
DI (2013) Increases in alcohol consumption in women and elderly
groups: evidence from an epidemiological study. BMC Public
Health 13:207

Hagenbeek FA, van Dongen J, Kluft C, Hankemeier T, Ligthart L,
Willemsen G, de Geus EJC, Vink JM, Bartels M, Boomsma DI (2018)
Adult aggressive behavior in humans and biomarkers: a focus on
lipids and methylation. J Pediatr Neonat Individ Med 7(2):e070204

Hamshere ML, Langley K, Martin J, Agha SS, Stergiakouli E, Anney
RJ, Buitelaar J, Faraone SV, Lesch KP, Neale BM, Franke B,
Sonuga-Barke E, Asherson P, Merwood A, Kuntsi J, Medland SE,
Ripke S, Steinhausen HC, Freitag C, Rief A, Renner TJ, Roma-
nos M, Romanos J, Warake A, Meyer J, Palmoson H, Vasquez
AA, Lambrechts-Rommelse N, Roeyers H, Biederman J, Doyle
AE, Hakonarson H, Rothenberger A, Banaschewski T, Oades
R, McDougall JJ, Kent L, Williams N, Owen MJ, Holmans P,
O’Donovan MC, Thapar A (2013) High loading of polygenic risk
for ADHD in children with comorbid aggression. Am J Psychiatry
170(8):909–916

Heath AC, Kessler RC, Neale MC, Eaves LJ, Kendler KS (1993) Test-
ing hypotheses about direction of causation using cross-sectional
family data. Behav Genet 23:29–50

Hinshaw SP (1992) Externalizing behavior problems and academic
underachievement in childhood and adolescence: causal relation-
ships and underlying mechanisms. Psychol Bull 111:127–155

Hinshaw SP, Owens EB, Sami N, Fargeon S (2006) Prospective fol-
low-up of girls with attention-deficit/hyperactivity disorder into
adolescence: Evidence for continuing cross-domain impairment.
J Consult Clin Psychol 74(3):489–499

Hurl YM (2015) Genetic and environmental etiology of the relation-
ship between childhood hyperactivity/inattention and conduct
problems in a South Korean twin sample. Twin Res Hum Genet
18(3):290–297

Ip HF et al (2019) Genetic association study of childhood aggres-
sion across raters, instruments and age. BioRxiv. https://doi.
org/10.1101/854927
Jensen PS, Martin D, Cantwell DP (1997) Comorbidity in ADHD: implications for research, practice, and DSM-V. J Am Acad Child Adolesc Psychiatry 36:1065–1079
Lee SS, Hinshaw SP (2006) Predictors of adolescent functioning in girls with attention deficit hyperactivity disorder (ADHD): the role of childhood ADHD, conduct problems, and peer status. J Clin Child Adolesc Psychol 35(3):356–368
Ligthart L, van Beijsterveldt CEM, Kevenaar ST, de Zeeuw E, van Bergen E, Bruins S, Pool R, Helmer Q, van Dongen J, Hottenga JJ, Van’t Ent D, Dolan CV, Davies GE, Ehli EA, Bartels M, Willemse G, de Geus EJC. Boomsma DI (2019) The Netherlands twin register: longitudinal research based on twin and twin-family designs. Twin Res Hum Genet 22(6):623–636
Lucaroni F, Cicciarella Modica D, Macino M, Palombi L, Abbondanzieri A, Agosti G, Biondi G, Morciano L, Vinci A (2019) Can risk be predicted? an umbrella systematic review of current risk prediction models for cardiovascular diseases, diabetes and hypertension. BMJ Open 9(12):e030234
Martin NG, Eaves LJ (1977) The genetical analysis of covariance structure. Heredity (Edinb) 38(1):79–95
Neale MC, Eaves LJ, Hewitt JK, Kendler KS (1994) Multiple regression with data collected from relatives: testing assumptions of the model. Multivar Behav Res 29:33–61
Neale MC, Hunter MD, Pritikin JN, Zahery M, Brick TR, Kirkpatrick RM, Estabrook R, Bates TC, Maes HH, Boker SM (2016) OpenMx 20: extended structural equation and statistical modeling. Psychometrika 81(2):535–549
Odintsova VV, Roetman PJ, Ip HF, Pool R, Van der Laan CM, Tona KD, Vermeiren RRJ, Boomsma DI (2019) Genomics of human aggression: current state of genome-wide studies and an automated systematic review tool. Psychiatr Genet 29(5):170–190
Ozaki K, Toyoda I, Iwama N, Kubo S, Ando J (2011) Using non-normal SEM to resolve the ACDE model in the classical twin design. Behav Genet 41(2):329–339
Porsch RM, Middeldorp CM, Cherny SS, Krapohl E, van Beijsterveldt CE, Loukola A, Korhonen T, Pulkkinen L, Corley R, Rhee S, Kaprio J, Rose RR, Hewitt JK, Sham P, Plomin R, Boomsma DI, Bartels M (2016) Longitudinal heritability of childhood aggression. Am J Med Genet B 171(5):697–707
Purcell S (2002) Variance components models for gene-environment interaction in twin analysis. Twin Res 5(6):554–571
Saad JF, Griffiths KR, Kohn MR, Clarke S, Williams LM, Korgaonkar MS (2017) Regional brain network organization distinguishes the combined and inattentive subtypes of attention deficit hyperactivity disorder. Neuroimage Clin 15:383–390
Saviouk V, Hottenga JJ, Slagboom EP, Distel MA, de Geus EJ, Willemse G, Boomsma DI (2011) ADHD in Dutch adults: heritability and linkage study. Am J Med Genet B 156B(3):352–362
Sokolova E, Groot P, Claassen T, van Hulzen KJ, Glennon JC, Franke B, Heskes T, Buitelaar J (2016) Statistical evidence suggests that inattention drives hyperactivity/impulsivity in attention deficit-hyperactivity disorder. PLoS ONE 11(10):e016512
van Beijsterveldt CE, Bartels M, Hudziak JJ, Boomsma DI (2003) Causes of stability of aggression from early childhood to adolescence: a longitudinal genetic analysis in Dutch twins. Behav Genet 33(5):591–605
van Beijsterveldt CE, Groen-Blokhuis M, Hottenga JJ, Franic S, Hudziak JJ, Lamb D, Huppermertz C, de Zeeuw E, Nivard M, Schutte N, Swagerman S, Glasner T, van Fulpen M, Brouwer C, Stroet T, Nowotny D, Ehli EA, Davies GE, Scheet P, Orlebeke JF, Kan KJ, Smit D, Dolan CV, Middeldorp CM, de Geus EJ, Bartels M, Boomsma DI (2013) The Young Netherlands Twin Register (YNTR): longitudinal twin and family studies in over 70,000 children. Twin Res Hum Genet 16(1):252–267
Vierikko E, Pulkkinen L, Kaprio J, Rose RJ (2004) Genetic and environmental influences on the relationship between aggression and hyperactivity-impulsivity as rated by teachers and parents. Twin Res 7(3):261–274
Vuoksimaa E, Rose RJ, Pulkkinen L, Palviainen T, Rimfeld K, Lundström S, Bartels M, van Beijsterveldt C, Hendriks A, de Zeeuw EL, Plomin R, Lichtenstein P, Boomsma DI, Kaprio J (2020) Higher aggression is related to poorer academic performance in compulsory education. J Child Psychol Psychiatry. https://doi.org/10.1111/jcpp.13273
Willcutt EG, Nigg JT, Pennington BF, Solanto MV, Rohde LA, Tannock R, Loo SK, Carlson CL, McBurnett K, Lahey BB (2012) Validity of DSM-IV attention deficit/hyperactivity disorder symptom dimensions and subtypes. J Abnorm Psychol 121(4):991–1010

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.