



Polygenic Risk for Major Depression Interacts with Parental Criticism in Predicting Adolescent Depressive Symptom Development

Stefanie A. Nelemans¹ · Marco Boks² · Bochao Lin³ · Tineke Oldehinkel⁴ · Pol van Lier⁵ · Susan Branje¹ · Wim Meeus^{1,6}

Received: 29 June 2020 / Accepted: 2 November 2020 / Published online: 23 November 2020
© The Author(s) 2020

Abstract

Research has focused more and more on the interplay between genetics and environment in predicting different forms of psychopathology, including depressive symptoms. While the polygenic nature of depressive symptoms is increasingly recognized, only few studies have applied a polygenic approach in gene-by-environment interaction ($G \times E$) studies. Furthermore, longitudinal $G \times E$ studies on developmental psychopathological properties of depression are scarce. Therefore, this 6-year longitudinal community study examined the interaction between genetic risk for major depression and a multi-informant longitudinal index of critical parenting in relation to depressive symptom development from early to late adolescence. The sample consisted of 327 Dutch adolescents of European descent (56% boys; $M_{\text{age } T_1} = 13.00$, $SD_{\text{age } T_1} = 0.44$). Polygenic risk for major depression was based on the Hyde et al. (*Nature Genetics*, 48, 1031–1036, 2016) meta-analysis and genetic sensitivity analyses were based on the 23andMe discovery dataset. Latent Growth Models suggested that polygenic risk score for major depression was associated with higher depressive symptoms across adolescence (significant main effect), particularly for those experiencing elevated levels of critical parenting (significant $G \times E$). These findings highlight how polygenic risk for major depression in combination with a general environmental factor impacts depressive symptom development from early to late adolescence.

Keywords Depressive symptoms · Adolescence · Longitudinal · Polygenic risk score (PRS) · Parenting · Gene-by-environment interaction ($G \times E$)

These authors jointly supervised this work: Susan Branje, Wim Meeus

✉ Stefanie A. Nelemans
s.a.nelemans@uu.nl

¹ Department of Youth and Family, Utrecht University, Utrecht, The Netherlands

² Department of Psychiatry, University Medical Center Utrecht, Utrecht, The Netherlands

³ Department of Translational Neuroscience, Brain Center Rudolf Magnus, Utrecht, The Netherlands

⁴ Department of Psychiatry, University Medical Center Groningen, Groningen, The Netherlands

⁵ Department of Developmental Psychology, VU University Amsterdam, Amsterdam, The Netherlands

⁶ Department of Developmental Psychology, Tilburg University, Tilburg, The Netherlands

Introduction

Major Depressive Disorder (MDD; American Psychiatric Association 2013) is a common (Kessler et al. 2007) and persistent disorder (Eaton et al. 2008) that is associated with a variety of comorbid mental health problems and impaired functioning in a wide range of domains (Maske et al. 2016). A clinical diagnosis of MDD represents the extreme end of a continuous distribution of symptom severity at the population level (Widiger and Samuel 2005). The same symptoms that define MDD are variable in the general population (Hankin et al. 2005; Liu 2016) and adolescence is a critical period for the development of such symptoms (Kessler et al. 2012). Consequently, research that focuses on the development of depressive symptoms in adolescence and identifies factors that affect this development is essential. While research has focused increasingly on the interplay between genetics and environment, studies in adolescence are scarce and few have used a longitudinal design to consider the developmental psychopathological properties of depression. Moreover, only

few studies have applied a polygenic approach in gene-by-environment interaction ($G \times E$) studies. Therefore, this longitudinal community study examined the interaction between polygenic risk for major depression and critical parenting in relation to depressive symptom development from early to late adolescence.

The etiology of depression is complex and results from both genetic and environmental factors (Garber and Rao 2014). Concerning genetic factors, twin and family studies have indicated that MDD is a moderately heritable trait (Sullivan et al. 2000). Despite robust evidence for heritability of MDD, unraveling the genetic architecture of complex traits such as MDD and identifying specific “vulnerability genes” has proven quite a challenge. Recent progress in molecular genetic analyses has led to the critical insight that the genetic basis for complex traits such as MDD is polygenic (Purcell et al. 2009), that is, resulting from the additive effect of many genetic variants (single-nucleotide polymorphisms; SNPs) with small effect sizes each. This has caused the field to move from investigating simple, single genetic markers (i.e., a so-called *candidate gene approach*) to more complex genetic indices based on multiple genes (i.e., a so-called *polygenic approach*; Belsky and Israel 2014; Wray et al. 2014). Moreover, whereas initial genetic main effect research was strongly hypothesis-driven and biologically based by focusing on genetic variation associated with specific biological functions (e.g., genetic variation associated with neurotransmitters dopamine and serotonin), research in the field has also become more hypothesis-free and data-driven (or data-inferred) in genome-wide association studies. Concerning the latter, a recent large-scale meta-analysis was among the first to identify 17 independent SNPs reaching genome-wide significance ($p < 5 \times 10^{-8}$) that were robustly associated with increased risk of self-reported major depression in adults of European descent (Hyde et al. 2016). Because major depression represents the extreme end of a continuous distribution of depressive symptoms at the population level (Widiger and Samuel 2005), this polygenic risk score for major depression may also be relevant for adolescent depressive symptom development. However, it is still unknown to this moment whether and how polygenic risk for major depression as identified among adults is associated with the development of depressive symptoms in adolescents from the general population. Hence, the first goal of the present study was to rely on findings from the Hyde et al. (2016) meta-analysis to examine the association between polygenic risk for major depression and depressive symptom development in an adolescent community sample. Because sex differences in depression have been found to emerge in adolescence (Hankin et al. 2015), it was also explored whether girls could be more genetically vulnerable

in a way that genetic risk for major depression is more strongly associated with adolescent depressive symptom development for girls compared to boys (Merikangas and Almasy 2020; in part, for example, through affected hormonal processes; Naninck et al. 2011).

In addition to genetic factors, environmental factors are assumed to play a prominent role in the development of depressive symptoms in adolescence (Garber and Rao 2014). An important environmental factor in adolescent depressive symptom development is rejecting and particularly aversive parenting, which has been modestly but systematically associated with higher levels of child and adolescent depressive symptoms across studies (for a meta-analysis, see McLeod et al. 2007). Parental criticism, which refers to negative comments expressed by parents to their child and reflects a non-supportive and critical emotional family climate, may be specifically important to consider in this respect. Historically, a critical family climate, which is at the core of Brown’s Expressed Emotion theory (Brown et al. 1972; further refined by Vaughn and Leff 1976), has been associated with relapse in depressed patients (Hooley 2007). More recently, parental criticism had received increased attention in association with the development of depressive symptoms among youth (Peris and Miklowitz 2015), also since the importance of (perceived) parental criticism on youth psychological adjustment has become central in other prominent theories, such as Parental Acceptance-Rejection Theory (PARTheory; Rohner et al. 2005). Importantly, studies have indeed shown that higher levels of parental criticism are associated with higher levels of adolescent depressive symptoms, both cross-sectionally and longitudinally (e.g., Nelemans et al. 2014). It has been suggested that adolescents may internalize parental criticism as self-criticism or a negative evaluation of the self and significant others, which in turn is associated with higher depressive symptoms (e.g., Bolton et al. 2009; Rohner et al. 2005).

However, not all adolescents exposed to parental criticism develop depressive symptoms to the same extent, which raises questions about individual differences in vulnerability to adverse environments. Importantly, research on gene-by-environment interactions ($G \times E$) builds on the assumption that people vary in the extent to which they are affected by environmental factors and that this sensitivity to the environment may be genetically predisposed (Pluess 2015). Historically, the two most prominent theories on environmental sensitivity are the Diathesis-Stress (or Dual-Risk) framework and the Differential Susceptibility framework (which are currently integrated in single overarching meta-framework of environmental sensitivity; Pluess 2015). On the one hand, in the Diathesis-Stress framework (Monroe and Simons 1991) environmental sensitivity is

seen primarily as vulnerability for developing problematic outcomes in individuals faced with environmental adversity. In the context of this study, the negative effect of parental criticism on adolescent depressive symptom development would be expected to be greatest for youth with stronger polygenic risk for major depression (i.e., genetically vulnerable youth). On the other hand, in the Differential Susceptibility framework (Belsky et al. 2007; Belsky and Pluess 2009) environmental sensitivity is seen as susceptibility with some being more and some less susceptible to both negative and positive environmental influences (*for better and for worse*).

Following pioneering $G \times E$ work on MDD (Caspi et al. 2003), several studies have supported interactions between individual genetic differences and exposure to adverse environments, such as chronic stress (e.g., Hammen et al. 2010), stressful or negative life events (e.g., Chen et al. 2013), and maladaptive parenting (e.g., Van Assche et al. 2017), in predicting depressive symptoms in youth (for a systematic review, see Dunn et al. 2011). Most of these studies were hypothesis-driven and investigated only a limited number of genetic variants in relevant biological pathways and findings generally suggested that adverse environments are particularly associated with higher depressive symptoms for youth with a genetic predisposition (in line with the Diathesis-Stress framework; Monroe and Simons 1991). However, also many hypothesis-driven $G \times E$ studies have failed to replicate these findings (Dunn et al. 2011; Van der Auwera et al. 2018). Critical reflections on these hypothesis-driven $G \times E$ studies have resulted in recommendations for future research to include a polygenic approach (Dick et al. 2015), given that complex traits such as major depression appear to result from the additive effect of many genetic variants with small effect sizes individually. The second goal of the present study was therefore to investigate whether the effect of parental criticism, as indicator of an adverse parental environment, on adolescent depressive symptom development was moderated by polygenic risk for major depression. Given the emergence of sex differences in depression in adolescence (Hankin et al. 2015) and some indications for $G \times E$ in relation to depressive symptoms for girls but not boys (Hammen et al. 2010), sex differences in $G \times E$ predicting adolescent depressive symptom development were also explored. Adolescent girls show stronger sensitivity to interpersonal stress, in a way that they experience negative interpersonal events, such as parental criticism, as more stressful than boys, and this has been associated with gender differences in adolescent depressive symptom development (Rudolph 2009). It was therefore explored whether the hypothesized stronger negative effect of parental criticism on adolescent depressive symptom development for genetically vulnerable youth would be stronger for girls compared to boys.

Current Study

The present 6-year longitudinal community study aimed to examine (1) whether data-inferred polygenic risk for major depression (as identified among adults of European descent in a recent meta-analysis; Hyde et al. 2016) was associated with adolescent depressive symptom development from early to late adolescence, and (2) whether data-inferred polygenic risk for major depression moderated the effect of parental criticism on adolescent depressive symptom development from early to late adolescence. Concerning the main effect of data-inferred polygenic risk on adolescent depressive symptom development, in line with a developmental psychopathological perspective (e.g., Cicchetti and Rogosch 2002) it was hypothesized that polygenic risk for major depression as identified among adults (Hyde et al. 2016) would already express itself in adolescence in the form of higher initial depressive symptoms at the start of adolescence and/or increasing depressive symptoms across adolescence. In addition, potential sex differences were explored in the association between genetic risk for major depression and adolescent depressive symptom development. Concerning the effect of $G \times E$ on adolescent depressive symptom development, in line with the Diathesis-Stress framework (Monroe and Simons 1991) it was hypothesized that parental criticism would be particularly associated with higher initial depressive symptoms at the start of adolescence and/or increasing depressive symptoms across adolescence for adolescents with higher polygenic risk for major depression. In this way, the negative effect of parental criticism on adolescent depressive symptom development would be expected to be greatest for genetically vulnerable youth. In addition, sex differences in $G \times E$ predicting adolescent depressive symptom development were explored. For the main analyses, the data-inferred polygenic risk score was based on the 17 genetic variants across the genome that were most strongly associated with major depression ($p < 5 \times 10^{-8}$) in a recent meta-analysis (Hyde et al. 2016). Moreover, this study included several important genetic sensitivity analyses. For these sensitivity analyses, information from the 23andMe discovery dataset was used to calculate 12 additional data-inferred polygenic risk scores, each based on more genetic variants with increasingly weaker associations with major depression but thereby better reflecting the total SNP-heritability¹ of major depression (for more information, please see the section on “Polygenic Risk Scores for Major Depression” in “Methods” below).

¹ SNP-heritability = the maximum proportion of variation in the phenotype that can be explained by additive effects of SNPs.

Methods

Participants

Participants in this 6-year longitudinal community study were 369 adolescents (56.6% boys; $M_{\text{age}} T_1 = 13.00$, $SD_{\text{age}} T_1 = 0.44$; 7.1% from low SES families based on parents' job level), who provided saliva samples to extract DNA, and their mothers ($M_{\text{age}} T_1 = 44.65$, $SD_{\text{age}} T_1 = 4.44$). Participants were part of the ongoing Research on Adolescent Development And Relationships (RADAR—young cohort; <https://doi.org/10.17026/dans-zrb-v5wp>) study and identified themselves as ethnic Dutch. All participants attended the first grade of secondary school at the start of the study. Following quality control of the genetic data, including exclusion of ethnic outliers, the study sample consisted of 327 adolescents (56.3% boys; $M_{\text{age}} T_1 = 13.00$, $SD_{\text{age}} T_1 = 0.44$; 6.5% from low SES families based on parents' job level) and their mothers ($M_{\text{age}} T_1 = 44.72$, $SD_{\text{age}} T_1 = 4.25$).² There were no significant differences between participants in the total genetic sample ($N = 369$) and the study sample with high-quality genetic data ($n = 327$) concerning sex, $\chi^2(1) = 0.16$, $p = 0.69$, SES, $\chi^2(1) = 1.62$, $p = 0.20$, age, $F(1, 367) = 0.25$, $p = 0.62$, adolescent depressive symptoms T_1 , $F(1, 363) = 0.74$, $p = 0.39$, adolescent-reported parental criticism T_1 , $F(1, 355) = 0.60$, $p = 0.44$, and mother-reported criticism T_1 , $F(1, 366) = 1.16$, $p = 0.28$.

Sample attrition was low over time, with 312 of the 327 adolescents (4.6% attrition from wave 1 to wave 6) and 309 of the 327 mothers (5.5% attrition from wave 1 to wave 6) still participating at the sixth annual measurement wave. There were no significant differences between adolescents participating at all six measurement waves and those dropping out of the study concerning sex, $\chi^2(1) = 0.59$, $p = 0.44$, SES, $\chi^2(1) = 1.10$, $p = 0.30$, age, $F(1, 325) = 0.01$, $p = 0.93$, adolescent depressive symptoms T_1 , $F(1, 321) = 1.43$, $p = 0.23$, adolescent-reported parental criticism T_1 , $F(1, 317) = 0.13$, $p = 0.72$, and mother-reported criticism T_1 , $F(1, 324) = 1.26$, $p = 0.26$. No genetic data were missing for the study sample ($n = 327$). Most missing data on the variables of interest were the result of dropout, with some occasional missings being due to unavailability of participants at a specific measurement wave or participants' decision not to complete certain parts of the questionnaire at a specific measurement wave. Specifically, missing data on adolescent depressive symptoms ranged between 1.22 and 4.89% across waves, missing data on adolescent-reported parental criticism ranged between 2.45 and 6.12% across

waves, and missing data on mother-reported criticism ranged between 0 and 5.81% across waves. Little's Missing Completely at Random (MCAR) test showed that data were missing at random, $\chi^2(452) = 475.09$, $p = 0.22$. Missing data were handled in *Mplus* with Full Information Maximum Likelihood (Muthén and Muthén 1998).

Procedure

Participants were recruited from randomly selected primary schools in the western and central regions of the Netherlands (Western Europe). All participants and their parents received a complete description of the study and provided active written informed consent before the start of the study ($N = 497$). For six successive years, adolescents and their mothers completed annual questionnaires during a home visit and received a small monetary compensation for every wave they completed the questionnaires. In addition, at the fifth measurement wave both parents and adolescents were again asked for active consent to collect saliva samples from adolescents to obtain DNA. In total, 369 adolescents provided consent to provide a saliva sample to extract DNA, for which these participants received an additional small monetary compensation. The RADAR study was approved by the board of the local research institute and by the Medical Ethical Committees of Utrecht University and VU University Amsterdam.

DNA Processing and Genetic Profiling

Genotyping of all participants was done with the Affymetrix 6.0 array (McCarroll et al. 2008) using DNA from saliva samples. The genotype calling was performed with Birdseed 2 algorithm (Korn et al. 2008; McCarroll et al. 2008) with the Affymetrix 3.3 APT software on all samples simultaneously. Samples were removed if the Affymetrix CQC < 0.40, the genotyping calling rate < 0.90, the heterozygosity F value was < 0.10 or > 0.10, or the DNA gender of the sample did not match the phenotype gender. Samples were also removed if the 10 genetic principal components indicated a Caucasian European (CEU) ethnic outlier after projection of the study samples on the 1000 Genomes reference sample. Single nucleotide polymorphisms (SNPs) were filtered using Plink 1.07 (Purcell et al. 2007) based on the following criteria: No or incorrect mapping on Build 37 HG19 of the human genome, inconsistent calls in plate control samples with an error rate > 1%, < 0.95 genotyping rate, minor allele frequencies (MAF) < 0.01, Hardy–Weinberg equilibrium $p < 0.000001$. After this quality control (QC), all SNPs were strand aligned to the 1000 Genomes phase 1 integrated release version 3 panel (Altshuler et al. 2012). Subsequently, the genotype data were phased using SHAPEIT2 (Delaneau et al. 2013). All the samples were genotyped by the Affymetrix 6.0 array, with a

² Participants were asked to complete the questionnaires around the same months every year for 6 successive years (January–March) and their age therefore increased by almost exactly 1 year every wave.

high variant calling rate ($> 0.96\%$; N of genotyped SNPs in the samples ranged from 650,005 to 674,172). Imputation was performed from the HRC and 1000 Genomes Project Phase 3 (v5) reference panels using a two-stage approach. Pre-imputation phasing and imputation of genotype platform specific SNPs was carried out using the MACH software (Li et al. 2010). Subsequently, imputation of the reference set was carried out with Minimac. There were 16,619,494 SNPs after imputation for further post-imputation quality control.

After imputation, SNPs that failed the following criteria were excluded: $MAF > 0.01$, MAF and the reference allele frequency difference < 0.20 , and Hardy–Weinberg equilibrium > 0.00001 , Mendelian error < 0.02 . This resulted in 10,837,431 remaining SNPs. Furthermore, SNPs with imputation values above 0.8 and a $MAF > 0.01$ in both the discovery dataset and the target dataset were extracted ($n = 7,585,001$) and clumped using PLINK 1.90 (Chang et al. 2015) in two rounds to remove high linkage disequilibrium (LD) regions and avoid LD inflating the polygenic risk scores: (1) physical distance threshold 250 kb and LD threshold R^2 of 0.5, and (2) physical distance threshold 5000 kb and LD threshold R^2 of 0.2 (McLaughlin et al. 2017; Pries et al. 2020). Lastly, complex LD regions were removed including the Major Histocompatibility (MHC) region. In the end, 571,857 LD pruned SNPs remained for the polygenic risk score calculation.

Polygenic Risk Scores for Major Depression

For the main analyses, this study used the genome-wide association study summary statistics (i.e., the beta-values, effective alleles, and p values) from the a recent meta-analysis in individuals of European descent (Hyde et al. 2016) to calculate a data-inferred polygenic risk score for MDD based on the exact same 17 independent SNPs that reached genome-wide significance ($p < 5 \times 10^{-8}$) in their joint analysis of the 23andMe discovery dataset, the Psychiatric Genomics Consortium dataset (including the Psychiatric Genomics Consortium MDD, bipolar disorder, SCZ1, SCZ1 + SWE, and the SCZ2 datasets), and the 23andMe replication dataset. Since this calculated polygenic risk score only included information on those 17 SNPs that reached genome-wide significance ($p < 5 \times 10^{-8}$) in this meta-analysis, this score reflected the additive genetic risk of only those few genetic variants that were found to be most strongly associated with major depression. However, the maximum proportion of variation in major depression that can be explained by additive effects of SNPs is captured by many more additional genetic variants that are less strongly associated with major depression than $p < 5 \times 10^{-8}$.

For the sensitivity analyses, the full genome-wide association study summary statistics were therefore requested

for the 23andMe discovery dataset³ (i.e., the beta-values, effective alleles, and p values) to calculate 12 additional data-inferred polygenic risk scores for major depression in this study. Specifically, each of these scores included information on more and more genetic variants with increasingly weaker associations with major depression and thereby better reflected the total SNP-heritability of major depression. The following different p -value thresholds for the additional data-inferred polygenic risk scores were used: 5×10^{-7} (48 SNPs), 5×10^{-6} (136 SNPs), 5×10^{-5} (410 SNPs), 5×10^{-4} (1,771 SNPs), 5×10^{-3} (9,428 SNPs), 0.01 (15,779 SNPs), 0.05 (55,013 SNPs), 0.10 (95,470 SNPs), 0.20 (166,109 SNPs), 0.30 (229,496 SNPs), 0.40 (288,036 SNPs) and 0.50 (342,808 SNPs). This data-driven approach in which polygenic risk scores are based on different p value thresholds is common practice and recommended (Belsky and Israel 2014; Choi et al. 2020), particularly in cases of highly polygenic traits for which most genetic effects are very small. In such cases, more lenient p -value thresholds and thus more genetically-inclusive polygenic risk scores may better capture the total genetic risk associated with a certain phenotype (SNP-heritability). Importantly, the p -value thresholds used to create polygenic risk scores are very different from $p < 0.05$ in null hypothesis significance testing: Whereas polygenic risk scores based on only those SNPs that reached genome-wide significance ($p < 5 \times 10^{-8}$) or other strict p -value thresholds (e.g., $p < 5 \times 10^{-5}$) capture the additive effect of the genetic variants most strongly associated with a certain phenotype, a p -value threshold of 1.00 is expected to capture the maximum proportion of SNP heritability in a certain phenotype. From this point forward, the term “genetically-selective” polygenic risk scores is used to refer to the data-inferred polygenic risk scores based on stricter p -value thresholds (e.g., $p < 5 \times 10^{-7}$, $p < 5 \times 10^{-5}$) and “genetically-inclusive” polygenic risk scores to refer to the data-inferred polygenic risk scores based on more lenient p value thresholds (e.g., $p < 0.01$, $p < 0.30$). All data-inferred polygenic risk scores were created using PLINK version 1.90. Just like in the

³ Full GWAS statistics could only be requested for the 23andMe discovery dataset (for more information, please see the paragraph on “Data access” in the Online Methods of Hyde et al. 2016) and were not available for the joint meta-analysis of the 23andMe discovery dataset, the Psychiatric Genomics Consortium dataset (including the Psychiatric Genomics Consortium MDD, bipolar disorder, SCZ1, SCZ1 + SWE, and the SCZ2 datasets), and the 23andMe replication dataset. Importantly, the 23andMe discovery dataset represented the largest sample size within the joint analysis (75,607 cases of depression and 231,747 controls) compared to the two other included datasets (9240 cases and 9519 controls for the Psychiatric Genomics Consortium dataset and 45,773 cases and 106,354 controls for the 23andMe replication dataset, respectively). Specifically, the 23andMe discovery dataset consists of 62% females with a European ancestry, with the majority (more than 87%) being 30 years or older (for more details, see Hyde et al. 2016).

recent meta-analysis (Hyde et al. 2016) this study built upon, all data-inferred polygenic risk scores were corrected for population stratification (i.e., the top five Dutch principal components resulting from PLINK's Principal Components Analysis,⁴ Abdellaoui et al. 2014, as recommended by Tucker et al. 2014), adolescents' age, and adolescents' sex before being included in any of the subsequent analyses.

Measures

Depressive symptoms

The Reynolds Adolescent Depression Scale, second edition (RADS-2; Reynolds 2000) was used to assess adolescent depressive symptoms. The RADS-2 is a self-report questionnaire that consists of 23 items, which are measured on a 4-point scale ranging from 1 (*almost never*) to 4 (*usually*). Sample items include "I am sad" and "I feel like crying." In this study, internal consistency for the total depression scale was found to be good across waves, with Cronbach's alpha ranging between 0.93 and 0.95. Previous studies have shown good psychometric properties for the RADS-2 among adolescents (Osman et al. 2010).

Parental criticism

The 5-item parental criticism subscale of the 38-item Level of Expressed Emotions Scale (LEE; Hale et al. 2007) was used to assess parental criticism as reported by adolescents and their mothers. Specifically, adolescents reported on their perceived criticism from both parents, while mothers reported on their own criticism towards their adolescent. Items were rated on a 4-point scale, ranging from 1 (*untrue*) to 4 (*true*). Sample items include "My parents are critical of me/I am critical of my child" and "My parents try to change me/I try to change my child". This study combined adolescent- and mother-reports of parental criticism across six successive years into a single multi-informant longitudinal average, with higher levels representing higher levels of parental criticism. Cronbach's alpha internal consistency for this multi-informant score was good across waves, ranging between 0.69 and 0.77. Previous studies have shown good psychometric properties for the 38-item LEE (Hale et al. 2007).

⁴ Sensitivity analyses suggested that polygenic risk scores controlled for all 20 Dutch principal components resulting from PLINK's Principal Components Analysis were strongly correlated with the polygenic risk scores controlled only for the top 5 Dutch principal components, $r_s \geq 0.958$, $p_s < 0.001$.

Statistical Analyses

Latent Growth Curve Models (LGCMs) in *Mplus* Version 8.3 (Muthén and Muthén 1998–2017) were constructed to capture initial levels of depressive symptoms at the first wave in early adolescence (i.e., the intercept) and change in these symptoms across 6 successive years (i.e., the linear and quadratic slopes). As a first step, the corrected polygenic risk scores for major depression were included as predictor of the intercept and slopes of adolescent depressive symptoms, whilst controlling for sex differences in adolescent depressive symptom development.⁵ As a second step, the multi-informant longitudinal index of parental criticism and the interaction between parental criticism and the polygenic risk scores for major depression were additionally included as predictors. Both the polygenic risk scores for major depression and multi-informant longitudinal index of parental criticism were standardized before creating the interaction terms. In this second step, the correlation between the polygenic risk scores for major depression and the multi-informant longitudinal index of parental criticism was also included to account for potential gene-environment correlation. Significant interactions between parental criticism and the polygenic risk scores for major depression were interpreted by visualizing predicted intercept or slope values of adolescent depressive symptoms at low (-1 SD) and high ($+1$ SD) levels of parental criticism and polygenic risk for major depression.

In all LGCMS, ML estimation with standard errors and chi-square robust to non-normality was used (i.e., MLR estimator; Muthén and Muthén 1998–2017). Model fit was assessed with the Comparative Fit Index (CFI), the Root Mean Squared Error of Approximation (RMSEA) and its 90% confidence interval (CI), and the Standardized Root Mean Square Residual (SRMR), using conventional standards (Hu and Bentler 1999; Kline 2005).

Results

Descriptive Statistics

A summary of the most relevant correlations among all study variables can be found in Table 1. Correlations between the corrected polygenic risk scores for major depression and adolescent depressive symptoms across time were in most instances significant but small in effect size, $r_s = 0.07$ – 0.20 , just like the effect size of correlations

⁵ Because sex differences in adolescent depressive symptoms have been robustly reported (e.g., Hankin et al. 2015), sex was included as a covariate in all analyses. Excluding sex as covariate from all analyses did not affect any of the main conclusions.

Table 1 Summary of correlations among the different polygenic risk scores and all study variables ($N = 327$)

Variable	1	2	3	4	5	6	7
1. Depressive symptoms T_1	–						
2. Depressive symptoms T_2	0.58***	–					
3. Depressive symptoms T_3	0.56***	0.66***	–				
4. Depressive symptoms T_4	0.55***	0.64***	0.71***	–			
5. Depressive symptoms T_5	0.51***	0.57***	0.64***	0.79***	–		
6. Depressive symptoms T_6	0.51***	0.53***	0.61***	0.70***	0.75***	–	
7. Parental criticism ^a	0.37***	0.39***	0.38***	0.43***	0.42***	0.37***	–
8. PRS MDD 5×10^{-8b}	0.15**	0.13*	0.16**	0.17**	0.13*	0.11*	0.05
9. PRS MDD 5×10^{-7}	0.10	0.18***	0.18**	0.19***	0.19***	0.14*	0.09
10. PRS MDD 5×10^{-6}	0.07	0.17**	0.16**	0.17**	0.17**	0.12*	0.06
11. PRS MDD 5×10^{-5}	0.11*	0.12*	0.15**	0.18***	0.20***	0.14*	0.07
12. PRS MDD 5×10^{-4}	0.12*	0.11*	0.13*	0.16**	0.16**	0.14**	0.06
13. PRS MDD 5×10^{-3}	0.10	0.15**	0.15**	0.13*	0.11*	0.13*	0.08
14. PRS MDD 0.01	0.10	0.14**	0.15**	0.13*	0.13*	0.16**	0.09
15. PRS MDD 0.05	0.13*	0.18**	0.19***	0.16**	0.15**	0.17**	0.15**
16. PRS MDD 0.10	0.13*	0.18**	0.18**	0.15**	0.14*	0.17**	0.14**
17. PRS MDD 0.20	0.14*	0.19***	0.18**	0.17**	0.16**	0.18**	0.18**
18. PRS MDD 0.30	0.14*	0.18***	0.18**	0.16**	0.15**	0.18**	0.19***
19. PRS MDD 0.40	0.15**	0.19***	0.19***	0.17**	0.16**	0.18**	0.21***
20. PRS MDD 0.50	0.15**	0.19***	0.19***	0.17**	0.16**	0.18**	0.20***

PRS MDD [...] = polygenic risk score for major depression, calculated based on the mentioned p -value threshold

^aAdolescent- and mother-reports of parental criticism across 6 successive years were combined into a single multi-informant longitudinal average, with higher levels representing higher levels of parental criticism

^bThe PRS MDD 5×10^{-8} score was based on the meta-analysis by Hyde et al. (2016) and corrected for adolescent age, sex, and population stratification. The other 12 polygenic risk scores for major depression with different p -value thresholds were based on the 23andMe summary statistics and also corrected for adolescent age, sex, and population stratification

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

between the corrected polygenic risk scores for major depression and parental criticism, $r_s = 0.05$ – 0.21 . In addition, correlations between and adolescent depressive symptoms across time were significant and medium in effect size, $r_s = 0.37$ – 0.43 . Means and standard deviations of all variables in the statistical analyses can be found in Table 2.

Adolescent Depressive Symptom Development

Fit of the LGCM of adolescent depressive symptoms without any predictors was good, $\chi^2(12) = 20.66$, RMSEA [90% CI] = 0.05 [0.00, 0.08], CFI = 0.99, SRMR = 0.03. In the total sample, adolescent depressive symptoms showed an initial slight decrease, $b_{\text{linear slope}} = -0.06$, $p < 0.001$, followed by a slight increase over time, $b_{\text{quadratic slope}} = 0.01$, $p < 0.001$ (see Fig. 1). In all main models, significant main effects of sex on adolescent depressive symptom development were found. Specifically, girls showed higher intercept levels of adolescent depressive symptoms, $\beta_s = 0.47$ – 0.55 , $p_s < 0.001$, and a less steep initial decrease in

depressive symptoms, $\beta_{\text{linear slope}} = 0.53$ – 0.58 , $p_s = 0.002$ – 0.005 , followed by a less steep increase in depressive symptoms over time, $\beta_{\text{quadratic slope}} = -0.49$ – -0.54 , $p_s = 0.006$ – 0.011 . This suggests higher mean levels and stronger stability in depressive symptoms across time for adolescent girls compared to adolescent boys (see Fig. 1).

Polygenic Risk for Major Depression and Adolescent Depressive Symptom Development

Fit of the models including the corrected polygenic risk scores for major depression and adolescent sex as predictors of the intercept and both linear and quadratic growth of adolescent depressive symptoms was good, $\chi^2(18) \leq 29.47$, RMSEA [90% CI] ≤ 0.04 [0.01, 0.07], CFI = 0.99, SRMR = 0.03. Importantly, in line with expectations, higher polygenic risk (i.e., stronger genetic risk for major depression; as identified by Hyde et al. 2016) was significantly associated with higher intercept levels of adolescent depressive symptoms, $\beta = 0.17$, $p = 0.019$. Moreover, the sensitivity analyses suggested replication of this finding across nearly

Table 2 Means and standard deviations of all variables included in the statistical analyses ($N = 327$)

	<i>M</i>	<i>SD</i>
Depressive symptoms T ₁	1.62	0.49
Depressive symptoms T ₂	1.49	0.49
Depressive symptoms T ₃	1.52	0.53
Depressive symptoms T ₄	1.55	0.54
Depressive symptoms T ₅	1.53	0.53
Depressive symptoms T ₆	1.59	0.56
Adolescent-reported parental criticism T ₁	1.64	0.47
Adolescent-reported parental criticism T ₂	1.67	0.51
Adolescent-reported parental criticism T ₃	1.76	0.57
Adolescent-reported parental criticism T ₄	1.80	0.56
Adolescent-reported parental criticism T ₅	1.77	0.57
Adolescent-reported parental criticism T ₆	1.77	0.57
Mother-reported parental criticism T ₁	1.57	0.39
Mother-reported parental criticism T ₂	1.54	0.38
Mother-reported parental criticism T ₃	1.52	0.38
Mother-reported parental criticism T ₄	1.54	0.38
Mother-reported parental criticism T ₅	1.54	0.39
Mother-reported parental criticism T ₆	1.52	0.38
Parental criticism ^a	1.63	0.32
PRS MDD 5×10^{-8b}	-8.75×10^{-6}	2.87×10^{-3}
PRS MDD 5×10^{-7}	-3.92×10^{-5}	2.38×10^{-3}
PRS MDD 5×10^{-6}	-1.56×10^{-5}	1.35×10^{-3}
PRS MDD 5×10^{-5}	-1.67×10^{-5}	7.15×10^{-4}
PRS MDD 5×10^{-4}	-4.88×10^{-6}	2.88×10^{-4}
PRS MDD 5×10^{-3}	-7.41×10^{-8}	1.14×10^{-4}
PRS MDD 0.01	6.84×10^{-7}	8.78×10^{-5}
PRS MDD 0.05	1.28×10^{-6}	4.41×10^{-5}
PRS MDD 0.10	1.12×10^{-6}	3.16×10^{-5}
PRS MDD 0.20	9.45×10^{-7}	2.20×10^{-5}
PRS MDD 0.30	7.19×10^{-7}	1.75×10^{-5}
PRS MDD 0.40	6.32×10^{-7}	1.48×10^{-5}
PRS MDD 0.50	5.52×10^{-7}	1.28×10^{-5}

PRS MDD [...] = polygenic risk score for major depression, calculated based on the mentioned p -value threshold

^aAdolescent- and mother-reports of parental criticism across six successive years were combined into a single multi-informant longitudinal average, with higher levels representing higher levels of parental criticism. For the analyses, this score was standardized to facilitate estimation and interpretation

^bThe PRS MDD 5×10^{-8} score was based on the meta-analysis by Hyde et al. (2016) and corrected for adolescent age, sex, and population stratification. The other 12 PRS MDD scores with different p -value thresholds were based on the 23andMe summary statistics and also corrected for adolescent age, sex, and population stratification. For the analyses, all PRS MDD were standardized to facilitate estimation and interpretation

all 12 additional polygenic risk scores; higher levels of polygenic risk for major depression were consistently significantly associated with higher intercept levels of adolescent depressive symptoms, $\beta_s = 0.09$ – 0.21 , $p_s = 0.001$ – 0.159 (see Table 3). Polygenic risk for major depression was not (consistently) significantly associated with changes in adolescent depressive symptoms over time (i.e., the linear or the quadratic slope).

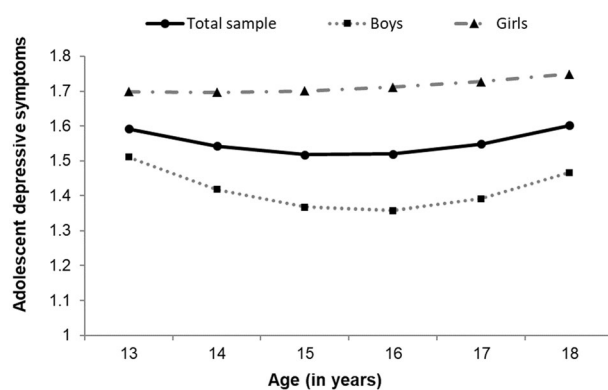


Fig. 1 Graphical representation of adolescent depressive symptom development from early to late adolescence across six successive waves (i.e., approximately ages 13–18 years) for the total sample ($N = 327$), as well as boys ($n = 184$) and girls ($n = 143$) separately

Fit of the models including the exploratory sex \times polygenic risk for major depression interactions was good, $\chi^2(21) \leq 33.00$, RMSEA [90% CI] ≤ 0.04 [0.01, 0.07], CFI = 0.99, SRMR = 0.03. The exploratory sex \times polygenic risk for major depression interactions were not consistently significantly associated with intercept levels of adolescent depressive symptoms, $\beta_s = 0.07$ – 0.24 , $p_s = 0.004$ – 0.390 , nor with the initial decrease in depressive symptoms, $\beta_{\text{linear slope}} = 0.11$ – 0.35 , $p_s = 0.007$ – 0.338 , or the following increase in depressive symptoms over time, $\beta_{\text{quadratic slope}} = -0.13$ – -0.35 , $p_s = 0.007$ – 0.311 (see Table 4).

Polygenic Risk for Major Depression \times Parental Criticism and Adolescent Depressive Symptom Development

Fit of the models including the interaction between the corrected polygenic risk scores for major depression and the multi-informant longitudinal index of parental criticism as predictors of the intercept and both linear and quadratic growth of adolescent depressive symptoms was good, $\chi^2(28) \leq 41.62$, RMSEA [90% CI] ≤ 0.04 [0.01, 0.06], CFI = 0.99–1.00, SRMR = 0.03–0.04. Importantly, in line with expectations, the interaction between the corrected polygenic risk score for major depression (i.e., genetic risk for major depression; as identified by Hyde et al. 2016) and parental criticism was significantly associated with higher intercept levels of adolescent depressive symptoms, $\beta = 0.15$, $p = 0.008$. A visualization of predicted intercept values of adolescent depressive symptoms at low (-1 SD) and high ($+1$ SD) levels of parental criticism and polygenic risk for major depression suggested that lower levels of parental criticism were associated with the lowest intercept levels of adolescent depressive symptoms, regardless of polygenic risk for major depression, whereas higher levels of parental criticism were associated with higher intercept

Table 3 Summary of standardized coefficients of polygenic risk for MDD predicting adolescent depressive symptom development, controlling for sex ($N = 327$)

Model	Intercept ^a	Linear slope	Quadratic slope
PRS MDD 5×10^{-8b}	$\beta = \mathbf{0.17}, p = \mathbf{0.019}$	$\beta = 0.08, p = 0.412$	$\beta = -0.10, p = 0.324$
PRS MDD 5×10^{-7}	$\beta = \mathbf{0.13}, p = \mathbf{0.038}$	$\beta = \mathbf{0.19}, p = \mathbf{0.022}$	$\beta = -\mathbf{0.18}, p = \mathbf{0.041}$
PRS MDD 5×10^{-6}	$\beta = 0.09, p = 0.159$	$\beta = \mathbf{0.21}, p = \mathbf{0.018}$	$\beta = -\mathbf{0.20}, p = \mathbf{0.034}$
PRS MDD 5×10^{-5}	$\beta = 0.12, p = 0.083$	$\beta = 0.15, p = 0.127$	$\beta = -0.12, p = 0.236$
PRS MDD 5×10^{-4}	$\beta = \mathbf{0.13}, p = \mathbf{0.043}$	$\beta = 0.08, p = 0.435$	$\beta = -0.05, p = 0.631$
PRS MDD 5×10^{-3}	$\beta = \mathbf{0.14}, p = \mathbf{0.017}$	$\beta = 0.08, p = 0.435$	$\beta = -0.07, p = 0.509$
PRS MDD 0.01	$\beta = \mathbf{0.16}, p = \mathbf{0.011}$	$\beta = 0.04, p = 0.696$	$\beta = -0.01, p = 0.946$
PRS MDD 0.05	$\beta = \mathbf{0.19}, p = \mathbf{0.002}$	$\beta = 0.07, p = 0.542$	$\beta = -0.05, p = 0.656$
PRS MDD 0.10	$\beta = \mathbf{0.19}, p = \mathbf{0.003}$	$\beta = 0.04, p = 0.711$	$\beta = -0.02, p = 0.842$
PRS MDD 0.20	$\beta = \mathbf{0.20}, p = \mathbf{0.002}$	$\beta = 0.05, p = 0.632$	$\beta = -0.04, p = 0.744$
PRS MDD 0.30	$\beta = \mathbf{0.20}, p = \mathbf{0.001}$	$\beta = 0.04, p = 0.701$	$\beta = -0.03, p = 0.812$
PRS MDD 0.40	$\beta = \mathbf{0.21}, p = \mathbf{0.001}$	$\beta = 0.06, p = 0.595$	$\beta = -0.05, p = 0.685$
PRS MDD 0.50	$\beta = \mathbf{0.21}, p = \mathbf{0.001}$	$\beta = 0.05, p = 0.654$	$\beta = -0.04, p = 0.729$

PRS MDD [...] = polygenic risk score for major depression, calculated based on the mentioned p -value threshold. Significant main effects ($p < 0.05$) are in bold

^aThe average age of participants at the intercept was 13.00 years old ($SD = 0.44$)

^bThe PRS MDD 5×10^{-8} score was based on the meta-analysis by Hyde et al. (2016) and corrected for adolescent age, sex, and population stratification. The other 12 polygenic risk scores for major depression with different p -value thresholds were based on the 23andMe summary statistics and also corrected for adolescent age, sex, and population stratification

Table 4 Summary of standardized coefficients of polygenic risk for MDD \times sex interactions predicting adolescent depressive symptom development ($N = 327$)

Model	Intercept ^a	Linear slope	Quadratic slope
PRS MDD 5×10^{-8b}	$\beta = 0.15, p = 0.123$	$\beta = 0.21, p = 0.081$	$\beta = -0.24, p = 0.052$
PRS MDD 5×10^{-7}	$\beta = \mathbf{0.20}, p = \mathbf{0.011}$	$\beta = 0.15, p = 0.181$	$\beta = -0.14, p = 0.264$
PRS MDD 5×10^{-6}	$\beta = \mathbf{0.24}, p = \mathbf{0.004}$	$\beta = 0.11, p = 0.338$	$\beta = -0.13, p = 0.311$
PRS MDD 5×10^{-5}	$\beta = \mathbf{0.24}, p = \mathbf{0.007}$	$\beta = 0.26, p = 0.055$	$\beta = -0.24, p = 0.089$
PRS MDD 5×10^{-4}	$\beta = 0.12, p = 0.185$	$\beta = \mathbf{0.35}, p = \mathbf{0.011}$	$\beta = -\mathbf{0.34}, p = \mathbf{0.015}$
PRS MDD 5×10^{-3}	$\beta = 0.07, p = 0.390$	$\beta = \mathbf{0.34}, p = \mathbf{0.008}$	$\beta = -\mathbf{0.35}, p = \mathbf{0.010}$
PRS MDD 0.01	$\beta = 0.09, p = 0.294$	$\beta = \mathbf{0.34}, p = \mathbf{0.007}$	$\beta = -\mathbf{0.35}, p = \mathbf{0.007}$
PRS MDD 0.05	$\beta = 0.15, p = 0.084$	$\beta = 0.25, p = 0.087$	$\beta = -0.28, p = 0.066$
PRS MDD 0.10	$\beta = 0.14, p = 0.107$	$\beta = 0.22, p = 0.135$	$\beta = -0.25, p = 0.104$
PRS MDD 0.20	$\beta = 0.15, p = 0.098$	$\beta = 0.23, p = 0.119$	$\beta = -0.27, p = 0.077$
PRS MDD 0.30	$\beta = 0.17, p = 0.053$	$\beta = 0.18, p = 0.224$	$\beta = -0.23, p = 0.148$
PRS MDD 0.40	$\beta = \mathbf{0.17}, p = \mathbf{0.040}$	$\beta = 0.19, p = 0.194$	$\beta = -0.23, p = 0.141$
PRS MDD 0.50	$\beta = 0.16, p = 0.053$	$\beta = 0.20, p = 0.176$	$\beta = -0.23, p = 0.128$

PRS MDD [...] = polygenic risk score for major depression, calculated based on the mentioned p -value threshold. Significant interactions ($p < 0.05$) are in bold

^aThe average age of participants at the intercept was 13.00 years old ($SD = 0.44$)

^bThe PRS MDD 5×10^{-8} score was based on the meta-analysis by Hyde et al. (2016) and corrected for adolescent age, sex, and population stratification. The other 12 polygenic risk scores for major depression with different p -value thresholds were based on the 23andMe summary statistics and also corrected for adolescent age, sex, and population stratification

levels of adolescent depressive symptoms, particularly for adolescents with higher polygenic risk for major depression (see Fig. 2). The sensitivity analyses suggested that this finding could not be replicated for most of the 12 additional polygenic risk scores, β s = -0.11 – -0.13 , p s = 0.019 – 0.891 (see Table 5).

The interaction between the corrected polygenic risk score for major depression (i.e., genetic risk for major depression, as identified by Hyde et al. 2016) and parental criticism was not significantly associated with changes in adolescent depressive symptoms over time (i.e., the linear or the quadratic slope). However, results from the sensitivity analyses suggested that

most of the 12 additional polygenic risk scores interacted significantly with parental criticism in predicting linear changes in adolescent depressive symptoms over time, β s = 0.11–0.25, p s = 0.007–0.184 (see Table 5). A visualization of predicted linear slope values of adolescent depressive symptoms at low (–1 SD) and high (+1 SD) levels of parental criticism polygenic risk for major depression suggested an interesting change in the pattern of interactions from the genetically-selective polygenic risk scores to the more

genetically-inclusive polygenic risk scores (see Fig. 3). Specifically, for the genetically-selective polygenic risk scores (i.e., $p < 5 \times 10^{-7}$ and $p < 5 \times 10^{-6}$) lower levels of parental criticism were associated with lower slope levels of adolescent depressive symptoms, regardless of polygenic risk for major depression, whereas higher levels of parental criticism were associated with higher slope levels for adolescents with higher polygenic risk for major depression but lower slope levels for adolescents with lower polygenic risk for major depression (see Fig. 3a, b). In contrast, for the more genetically-inclusive polygenic risk scores (i.e., $p < 0.05$ – $p < 0.50$) adolescents with high polygenic risk for major depression showed the lowest slope levels of adolescent depressive symptoms at low levels of parental criticism, but the highest slope levels at high levels of parental criticism (see Fig. 3c, d). Adolescents with low polygenic risk for major depression showed higher slope levels of adolescent depressive symptoms at low levels of parental criticism and lower slope levels at high levels of parental criticism (see Fig. 3c, d). Thus, for the more genetically-inclusive polygenic risk scores findings suggested a cross-over interaction.

The models with the exploratory adolescent sex \times polygenic risk for major depression \times parental criticism interactions failed to achieve acceptable fit to the data, and resulted in estimation problems in attempts to improve fit, and the model complexity exceeded this study’s sample size. Therefore, the results of these exploratory analyses were not reported.

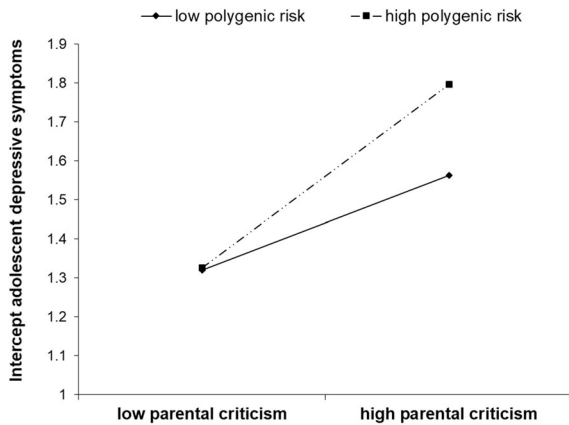


Fig. 2 Graphical representation of the significant interaction between the corrected polygenic risk score for major depression (based on Hyde et al. 2016) and the multi-informant longitudinal index of parental criticism on intercept levels of adolescent depressive symptoms

Table 5 Summary of standardized coefficients of polygenic risk for MDD \times parental criticism interactions predicting adolescent depressive symptom development, controlling for sex ($N = 327$)

Model	Intercept ^a	Linear slope	Quadratic slope
PRS MDD 5×10^{-8b}	$\beta = \mathbf{0.15}, p = \mathbf{0.008}$	$\beta = 0.11, p = 0.184$	$\beta = -0.10, p = 0.241$
PRS MDD 5×10^{-7}	$\beta = \mathbf{0.13}, p = \mathbf{0.019}$	$\beta = \mathbf{0.20}, p = \mathbf{0.015}$	$\beta = -\mathbf{0.19}, p = \mathbf{0.037}$
PRS MDD 5×10^{-6}	$\beta = \mathbf{0.01}, p = \mathbf{0.829}$	$\beta = \mathbf{0.24}, p = \mathbf{0.007}$	$\beta = -\mathbf{0.22}, p = \mathbf{0.015}$
PRS MDD 5×10^{-5}	$\beta = 0.01, p = 0.891$	$\beta = 0.20, p = 0.051$	$\beta = -0.14, p = 0.147$
PRS MDD 5×10^{-4}	$\beta = -0.02, p = 0.730$	$\beta = 0.15, p = 0.145$	$\beta = -0.08, p = 0.436$
PRS MDD 5×10^{-3}	$\beta = -0.05, p = 0.440$	$\beta = \mathbf{0.17}, p = \mathbf{0.048}$	$\beta = -0.15, p = 0.087$
PRS MDD 0.01	$\beta = -0.05, p = 0.452$	$\beta = 0.13, p = 0.131$	$\beta = -0.09, p = 0.302$
PRS MDD 0.05	$\beta = -0.06, p = 0.301$	$\beta = \mathbf{0.22}, p = \mathbf{0.016}$	$\beta = -0.17, p = 0.081$
PRS MDD 0.10	$\beta = -0.09, p = 0.144$	$\beta = \mathbf{0.25}, p = \mathbf{0.012}$	$\beta = -0.19, p = 0.066$
PRS MDD 0.20	$\beta = -0.11, p = 0.084$	$\beta = \mathbf{0.24}, p = \mathbf{0.019}$	$\beta = -0.18, p = 0.085$
PRS MDD 0.30	$\beta = -0.10, p = 0.102$	$\beta = \mathbf{0.24}, p = \mathbf{0.018}$	$\beta = -0.19, p = 0.080$
PRS MDD 0.40	$\beta = -0.10, p = 0.107$	$\beta = \mathbf{0.24}, p = \mathbf{0.022}$	$\beta = -0.18, p = 0.092$
PRS MDD 0.50	$\beta = -0.11, p = 0.079$	$\beta = \mathbf{0.25}, p = \mathbf{0.014}$	$\beta = -0.20, p = 0.064$

PRS MDD [...] = polygenic risk score for major depression, calculated based on the mentioned p -value threshold. Significant interactions ($p < 0.05$) are in bold

^aThe average age of participants at the intercept was 13.00 years old ($SD = 0.44$)

^bThe PRS MDD 5×10^{-8} score was based on the meta-analysis by Hyde et al. (2016) and corrected for adolescent age, sex, and population stratification. The other 12 polygenic risk scores for major depression with different p -value thresholds were based on the 23andMe summary statistics and also corrected for adolescent age, sex, and population stratification

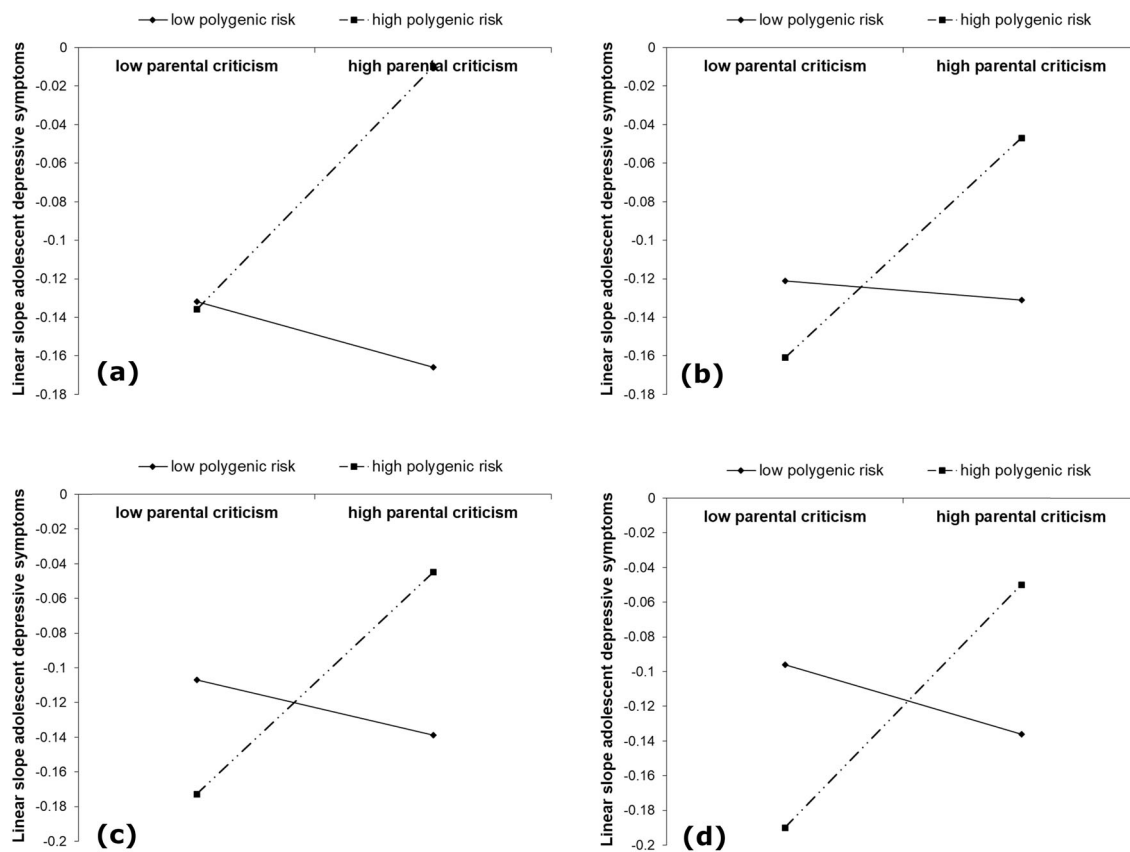


Fig. 3 Graphical representation of the significant interactions in the sensitivity analyses between the corrected polygenic risk scores for major depression and the multi-informant longitudinal index of

parental criticism on linear slope levels of adolescent depressive symptoms, for p -value thresholds (a) $p < 5 \times 10^{-7}$, (b) $p < 5 \times 10^{-3}$, (c) $p < 0.05$, and (d) $p < 0.50$

Discussion

Adolescence is a critical period for the development of depressive symptoms (Kessler et al. 2012) and research that focuses on identifying factors that affect this development is essential. The etiology of depression is complex (Garber and Rao 2014) and research has focused increasingly on how the interplay between genetics and environment ($G \times E$) are associated with depression. Yet, $G \times E$ studies on depression are scarce in adolescence, particularly longitudinal studies that focus on depressive symptom development and that apply a polygenic approach. Therefore, this study aimed to examine (1) whether polygenic risk for major depression (as identified among adults of European descent in a recent meta-analysis; Hyde et al. 2016) was associated with adolescent depressive symptom development from early to late adolescence, and (2) whether polygenic risk for major depression moderated the effect of parental criticism on adolescent depressive symptom development from early to late adolescence.

The findings from the present longitudinal community study suggested that polygenic risk for major depression, as

identified in a recent meta-analysis (Hyde et al. 2016), was associated with higher mean levels of depressive symptoms from early to late adolescence. Moreover, polygenic risk for major depression significantly moderated the association between parental criticism and mean levels of depressive symptoms from early to late adolescence. Specifically, as hypothesized, the highest mean levels of depressive symptoms across adolescence were found for those adolescents experiencing higher levels of parental criticism and with higher polygenic risk for major depression (i.e., genetically vulnerable individuals). Polygenic risk for major depression was not significantly associated with changes in adolescent depressive symptoms over time. Furthermore, sensitivity analyses suggested that that polygenic risk for major depression based on different p value thresholds was consistently associated with higher mean levels of depressive symptoms from early to late adolescence (see Table 3), suggesting robustness of the findings. Consistent with the main findings, polygenic risk for major depression did not consistently moderate the association between parental criticism and *mean levels* of depressive symptoms across adolescence in the sensitivity analyses. However, in contrast

to the main analyses polygenic risk for major depression did consistently moderate the association between parental criticism and *change* in depressive symptoms across adolescence in the sensitivity analyses (see Table 5). Moreover, in these analyses the pattern of the interactions appeared to change from the genetically-selective polygenic risk scores to the more genetically-inclusive polygenic risk scores (see Fig. 3). In sum, findings in this study suggested that polygenic risk for major depression was associated with higher depressive symptoms from early to late adolescence, particularly for those adolescents experiencing high levels of critical parenting. In addition, the findings highlight that different indices of genetic risk, for example in the present study captured by different *p*-value thresholds, may show different G × E interaction patterns.

The main effect of polygenic risk for major depression in adulthood on adolescent depressive symptom development showed that information on polygenic risk for major depression, as identified among adults in a large-scale GWAS (Hyde et al. 2016), is also highly relevant for predicting depressive symptom development from early to late adolescence in the general population. This is in line with the developmental psychopathology perspective (Cicchetti and Rogosch 2002) and a dimensional view on psychopathology (e.g., Hankin et al. 2005), in which major depression represents the extreme end of a continuous distribution of symptom severity at the population level. The finding that adolescents with higher polygenic risk for major depression showed higher mean (i.e., intercept) levels of depressive symptoms across adolescence suggests that genetic risk for major depression in adulthood already expresses itself in the form of heightened levels of depressive symptoms throughout adolescence. Polygenic risk for major depression was not significantly associated with changes in adolescent depressive symptoms over time on top of higher mean levels of depressive symptoms. This pattern of findings was replicated in the sensitivity analyses. Also, there were no consistent sex differences in the association between polygenic risk for major depression and mean levels or changes of depressive symptoms across adolescence. Whereas some have suggested that the emergence of sex differences in depressive symptoms in adolescence (Hankin et al. 2015) may be explained by higher vulnerability to genetic risk for adolescent girls compared to boys (Scourfield et al. 2003), no consistent evidence for such sex differences were found in this study.

In addition to the significant main effect of polygenic risk for major depression on mean levels of depressive symptoms across adolescence, in the main analyses genetic risk for major depression was also found to moderate the association between parental criticism and mean levels of depressive symptoms across adolescence (i.e., G × E). As hypothesized, parental criticism was particularly associated with higher

mean levels of depressive symptoms across adolescence for those adolescents with higher polygenic risk for major depression. At the same time, lower levels of parental criticism were associated with lower mean levels of adolescent depressive symptoms, regardless of polygenic risk for major depression (see Fig. 2). In line with the Diathesis–Stress framework (Monroe and Simons 1991), polygenic risk for major depression, captured by only those genetic variants that were found to reach genome-wide significance ($p < 5 \times 10^{-8}$) among adults in a recent meta-analysis (Hyde et al. 2016), thus appeared to distinguish adolescents based on vulnerability versus resilience. Specifically, the highest depressive symptoms across adolescence were found for those adolescents with both high levels of critical parenting and high polygenic risk for major depression (i.e., genetically vulnerable individuals), whereas mean levels of depressive symptoms across adolescence were not affected by critical parenting for adolescents with low polygenic risk for major depression (i.e., resilient individuals). In the sensitivity analyses, this significant G × E finding on mean levels of adolescent depressive symptoms was only replicated for the most genetically-selective polygenic risk score (i.e., $p < 5 \times 10^{-7}$), but not any of the other additional polygenic risk scores (see Table 5).

Whereas no significant G × E interaction was found in the prediction of changes in adolescent depressive symptoms over time in the main analyses, in the sensitivity analyses polygenic risk for major depression did quite consistently moderate the association between parental criticism and change in depressive symptoms across adolescence. Moreover, the pattern of the interactions appeared to change from the genetically-selective polygenic risk scores to the more genetically-inclusive polygenic risk scores (see Fig. 3). Specifically, for the genetically-selective polygenic risk scores (e.g., $p < 5 \times 10^{-7}$ and $p < 5 \times 10^{-5}$) low levels of parental criticism were associated with a (normative) decrease in adolescent depressive symptoms, regardless of genetic risk for major depression, whereas higher levels of parental criticism were associated with a weaker decrease (or rather, stability) in adolescent depressive symptoms for adolescents with higher compared to lower genetic risk for major depression. In contrast, for the more genetically-inclusive polygenic risk scores (i.e., $p < 0.05$ – 0.50) low levels of parental criticism were associated with a stronger decrease in adolescent depressive symptoms for adolescents with higher polygenic scores compared to adolescents with lower polygenic scores, whereas higher levels of parental criticism were associated with a weaker decrease (or rather, stability) in adolescent depressive symptoms for adolescents with higher polygenic scores compared to adolescents with lower polygenic scores (i.e., a cross-over interaction). Thereby, the interaction patterns appeared to change from being in line with the Diathesis–Stress framework (Monroe

and Simons 1991) for the more genetically-selective polygenic risk scores, in line with the $G \times E$ finding on mean levels of adolescent depressive symptoms in the main analyses, to being in line with the Differential Susceptibility framework (Belsky et al. 2007; Belsky and Pluess 2009) for the more genetically-inclusive thresholds, although no formal statistical tests were conducted to distinguish the interaction effects in terms of Diathesis–Stress versus Differential Susceptibility (see Roisman et al. 2012).

In terms of “substantive meaning”, genetically-selective polygenic risk scores are based on a few SNPs that are strongly significantly associated with the specific phenotype (e.g., $p < 5 \times 10^{-8}$). So, these polygenic risk scores likely represent genetic risk scores for a specific phenotype, such as (self-reported) major depression in this study, and appear to distinguish individuals based on environmental vulnerability versus resilience (in line with a Diathesis–Stress framework; Monroe and Simons 1991). In contrast, genetically-inclusive polygenic risk scores are based on thousands of SNPs across the genome with increasingly weaker associations with the specific phenotype, but thereby better reflecting the total SNP-heritability of major depression. In this study, these polygenic risk scores appeared to distinguish individuals based on environmental plasticity with the highest depressive symptoms for those adolescents with high polygenic risk at high levels of critical parenting, but also the lowest depressive symptoms at low levels of critical parenting (in line with the Differential Susceptibility framework; Belsky et al. 2007; Belsky and Pluess 2009). Importantly, past $G \times E$ interaction research on (adolescent) depressive symptoms has found evidence for findings in line with both Diathesis–Stress and Differential Susceptibility, and the present study highlights that different $G \times E$ interaction patterns may be found depending on the genetic index that is created (e.g., the applied p -value threshold for calculating a polygenic risk score).

Although it was not the main focus of this study, it is worth noting that correlations between the environmental factor parental criticism and the polygenic risk scores for MDD appeared to increase in strength with more inclusive p -value thresholds, particularly for those thresholds higher than 0.01 (see descriptive statistics in Table 1). As the more genetically-inclusive polygenic risk scores include more SNPs across the genome that are less strongly associated with the specific phenotype of interest, (self-reported) major depression in this study, these scores may include SNPs that are not uniquely associated with the phenotype of interest but may be associated with other phenotypes as well. Combined with the observation of stronger correlations between parental criticism and the more genetically-inclusive polygenic risk scores for MDD, one might question that these scores somehow reflect gene–environment correlation (rGE) associated with MDD either in itself or in

the $G \times E$ interactions. However, such rGE is not a likely explanation. First, polygenic risk scores with more inclusive p value thresholds more closely approximate the total SNP-heritability of the phenotype of interest, that is, the proportion of variation in a phenotype that can be explained by additive effects of observed commonly-occurring genetic variants or SNPs (h^2_{SNP}). While more inclusive p -value thresholds might be more likely to include some “environmentally sensitive” SNPs that may confound the polygenic risk scores’ reflection of the total heritability of a phenotype (h^2), this is not the same as capturing rGE. Second, the $G \times E$ analyses explicitly included the correlation between the polygenic risk scores and the environmental factor. Thereby, any potential environmental “confounding” in the $G \times E$ interactions is corrected for in these analyses and the polygenic risk scores’ explained variance would “purely” reflect the proportion of SNP heritability defined by the p -value threshold. It may, however, be an interesting direction for future research to systematically investigate whether more genetically-inclusive polygenic risk scores for MDD are more strongly associated with different relevant environmental factors compared to more genetically-selective polygenic risk scores and whether this appears to be the case across other phenotypes as well, to examine whether this phenomenon can be more widely observed than this study.

Implications

The fundamental knowledge resulting from this study on $G \times E$ interactions has some interesting implications, although cautious interpretation is warranted considering issues such as sample size and generalization (which is further elaborated upon below). First, it adds to the etiological literature (Garber and Rao 2014) on individual differences in (a) vulnerability to develop depressive symptoms (i.e., significant genetic and environmental main effects), and (b) vulnerability/sensitivity to environmental exposure related to depressive symptom development (i.e., significant $G \times E$ interactions). Another finding of note in this respect is the relevance of parenting as a “general” measure of environmental exposure. Whereas past research has often focused on “high impact” environmental factors, such as child maltreatment or abuse or traumatic/stressful life events (Dunn et al. 2011), this study found that normal variation in parenting experienced by adolescents from the normal population also has the potential to interact with (i.e., exaggerate or attenuate) genetic risk in predicting depressive symptom development across adolescence. So, in addition to a focus on such “high impact” environmental factors in clinical practice, normal variation in environmental experiences (such as parenting) may deserve explicit attention.

Second, this knowledge is relevant in the context of personalized medicine and therapygenetics (see discussion in e.g., Lester and Eley 2013). Therapygenetics refers to the prediction of psychological therapy outcomes from genetic markers. In this context, knowledge on adolescents' genetic susceptibility to the environment could potentially be used to predict their treatment response. Whereas genetically susceptible adolescents could potentially benefit from a brief and mild intervention, less susceptible adolescents might need a more intensive and comprehensive intervention to reduce depressive symptoms. Alternatively, genetic variation might differentially predict response to specific interventions, such as Cognitive Behavioral Therapy versus pharmacological treatment (e.g., Selective Serotonin Reuptake Inhibitors [SSRIs]). Within the broader context of personalized medicine, a fuller and more nuanced understanding of both genetic vulnerability and genetic sensitivity to the environment—or the developmental psychopathology of different depression more generally—might eventually aid in decisions on which treatment to select to maximize the chance of recovery for a particular individual. Even though personalized medicine concerning psychopathological treatment and therapygenetics are still in their infancy, in the future they may prove to be of great importance in guiding treatment selection and improving treatment effectiveness and outcome which, in turn, alleviates psychopathological symptoms and associated impaired functioning in many domains (Beevers and McGeary 2012).

Strengths, Limitations, and Directions for Future Research

This study has several strengths. First, the main analyses relied on information from a recent meta-analysis (Hyde et al. 2016), which includes the first robust findings concerning polygenic risk for major depression (identified through self-report questionnaires) in individuals of European descent based on GWAS. Moreover, this information was complemented with information from the large 23andMe dataset to address issues of replication, sensitivity, and robustness within this study's unique multi-informant longitudinal community sample. Second, by interacting polygenic risk for major depression with a multi-informant longitudinal index of critical parenting, this study followed recent recommendations for $G \times E$ studies to include polygenic approaches (Belsky and Israel 2014; Dick et al. 2015) and to increase the assessment quality of environmental exposure to increase power (Wong et al. 2003). Third and final, this study's 6-year longitudinal design captures an extended period from early to late adolescence and thereby offers critical information on potential predictors of depressive symptom development in a developmentally vulnerable period (Garber and Rao 2014), which is scarce as

most $G \times E$ research has been cross-sectional in nature and conducted in adult samples. Thereby, this study's data-inferred polygenic developmental psychopathological perspective on the study of genetic main effects and $G \times E$ interaction effects on adolescent depression addresses several gaps in the literature.

At the same time, this study should be considered in the light of some limitations, which may provide potential alternative explanations for adolescent depressive symptom development within a multifactorial developmental psychopathology framework (e.g., equifinality; Cicchetti and Rogosch 2002) as well as directions for future research. First, while within the current sample the results were replicated across several subsequent polygenic risk score p value thresholds, suggesting robustness of findings in this sample, these analyses do not preclude the need for independent replications as with all $G \times E$ analyses. Also, while this study represents a unique polygenic, longitudinal $G \times E$ interaction approach on adolescent depressive symptom development that provides important preliminary insights, cautious interpretation and replication of findings in much larger samples is warranted given the relatively small sample size. On a related note, caution should be taken in generalizing the findings from this study to other populations and situations. The present sample consists of a relatively well-functioning community sample of adolescents with a relatively homogeneous ethnic background from one country in Europe and it is unclear whether the present results can be extended to adolescents from other regions of the world, who have a more diverse socio-economic and ethnic background, and who are more diverse in functioning. Furthermore, future research should expand on this study focusing on other relevant parenting, or more broadly, other relevant environmental, variables. Also, testing the exploratory adolescent sex \times polygenic risk for major depression \times parental criticism interactions was unfortunately not possible in this study, for which larger longitudinal samples are required. It has also been suggested that genetic risk for major depression may interact in sex-specific ways with "interpersonal stressors" other than those in the parental environment, such as peer stressors, in predicting depressive symptom development (Oldehinkel and Bouma 2011; Shih et al. 2006).

Second, PRSs for major depression are purely "descriptive" in a way that they only represent a summary measure of genetic risk for major depression by the number of risk alleles associated with major depression as identified in GWAS and do not inform in any way on potential biological mechanisms underlying genetic risk. Thereby, all significant G and $G \times E$ findings are mere statistical results that are in need of an explanation in terms of their underlying (biological) mechanisms. On a related note, the field is currently rapidly evolving with new genome-wide

association meta-analyses for different phenotypes being published more commonly, including major depression (e.g., Wray et al. 2018). This study has specifically relied on a recent meta-analysis (Hyde et al. 2016) and the 23andMe dataset for the polygenic risk score calculations, because they relied on self-reports for identification of major depression case status. Since this study also relied on self-reports of depressive symptoms, this meta-analysis (Hyde et al. 2016) and the 23andMe dataset more closely match this sample than other studies, which may enhance the relevance of these genetic studies to the present study (although strong positive genetic correlations have been found between the 23andMe dataset and carefully curated clinical samples; Hyde et al. 2016; Wray et al. 2018).

Third and final, this study incorporated a longitudinal multi-informant average of critical parenting from early to late adolescence because of reasons of model complexity. Specifically, with 32 free parameters estimated in the current $G \times E$ analyses and a sample size of 327, the analyses were at the recommended number of cases to the number of free parameters ratio of 10:1 (Kline 2005, p. 111). Including a more complex representation of parental criticism in the $G \times E$ analyses would increase the number of free parameters and thereby bring the ratio cases/parameter ratio under the recommended 10:1, implying that model complexity would exceed the study's sample size. An important challenge for future research would be to incorporate longitudinal assessments of the environmental factor included in $G \times E$, in addition to longitudinal assessments of the phenotype of interest, because environmental factors such as parenting are dynamic in nature and may thereby show change, as well as differential effects, over time. Larger sample sizes are needed to, for example, include a latent longitudinal multi-informant index of parenting, in which each wave of each informant would be differentially weighted in contributing to the overall latent construct, or a Latent Growth Curve Model of (multi-informant) parenting, in which initial levels and change over time of the environmental factor are associated with those of the phenotype of interest and these associations would be moderated by genetic risk. Also, although the correlation between the genetic risk scores and critical parenting was included in all models to account for potential gene-environment correlations, potential intergenerational transmission of genetic risk must be acknowledged. This could either be direct transmission of genetic variation associated with depressive symptoms from parents to their adolescents or indirect transmission, as critical parenting can be rooted in depressive (or other psychiatric) symptoms in parents and thereby predict adolescents' depressive symptom development (i.e., mediation processes from G to E rather than moderation of $G \times E$; Harold et al. 2011; McAdams et al. 2014).

Conclusion

While research has focused increasingly on the interplay between genetics and environment ($G \times E$), studies in adolescence are scarce and few have used a longitudinal design to consider the developmental psychopathological properties of depression. Moreover, only few studies have applied a polygenic approach in gene-by-environment interaction studies. Therefore, this longitudinal community study examined the interaction between polygenic risk for major depression and critical parenting in relation to depressive symptom development from early to late adolescence. First, this study showed that polygenic risk for major depression is robustly associated with higher depressive symptoms across adolescence. Second, polygenic risk for major depression interacted with critical parenting in predicting depressive symptom development across adolescence, with the highest levels of depressive symptoms across adolescence for high polygenic risk combined with high levels of critical parenting (in line with the Diathesis–Stress framework; Monroe and Simons 1991). Overall, the results highlight how polygenic risk for major depression in combination with a general environmental factor impacts adolescent depressive symptoms from early to late adolescence. In addition, this study illustrates that different indices of genetic risk, e.g., by relying on different p value thresholds for their calculation, may show different $G \times E$ interaction patterns.

Acknowledgements The authors would like to thank all schools and families for their participation.

Authors' Contributions SN conceived of the study, participated in the data collection and coordination, performed the statistical analyses, interpreted the results, and drafted the manuscript; MB participated in the study design and coordination, was involved in the interpretation of the results, helped to draft the manuscript, and critically revised the manuscript; BL performed the statistical analyses, was involved in the interpretation of the results, helped to draft the manuscript, and critically revised the manuscript; TO participated in the study design and critically revised the manuscript; PvL, SB, and WM designed the larger project, participated in its design and coordination, were involved in the interpretation of the results, and critically revised the manuscript. All authors read and approved the final manuscript.

Funding Data of the RADAR (Research on Adolescent Development And Relationships) Young study were used (10.17026/dans-zrb-v5wp). RADAR has been financially supported by main grants from the Netherlands Organisation for Scientific Research (GB-MAGW 480-03-005, GB-MAGW 480-08-006), Stichting Achmea Slachtoffer en Samenleving (SASS), the Netherlands Organisation for Scientific Research to the Consortium Individual Development (CID; 024.001.003), and various other grants from the Netherlands Organisation for Scientific Research, VU University Amsterdam, and Utrecht University.

Data Sharing and Declaration The datasets generated and/or analyzed during the current study are available in the Data Archiving and Networked Services (DANS) repository, 10.17026/dans-zrb-v5wp.

Compliance with Ethical Standards

Conflict of Interest The authors declare 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 study was approved by the board of the local research institute and the Medical Ethical Committees of the Utrecht Medical Center and the VU Medical Center, the Netherlands.

Informed Consent Written informed consent was obtained from all individual participants (and their parents) included in the study.

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

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, 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 license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license 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 license, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Abdellaoui, A., Verweij, K. J. H., & Zietsch, B. P. (2014). No evidence for genetic assortative mating beyond that due to population stratification. *Proceedings of the National Academy of Sciences of the United States of America*, *111*, e4137.
- Altshuler, D. M., Durbin, R. M., Abecasis, G. R., Chakravarti, A., Clark, A. G., Donnelly, P., & Lachance, P. (2012). An integrated map of genetic variation from 1,092 human genomes. *Nature*, *491*, 56–65.
- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders*. 5th edn. Arlington, VA: American Psychiatric Publishing.
- Beevers, C. G., & McGeary, J. E. (2012). Therapygenetics: moving towards personalized psychotherapy treatment. *Trends in Cognitive Sciences*, *16*, 11–12.
- Belsky, J., Bakermans-Kranenburg, M. J., & van Ijzendoorn, M. H. (2007). For better and for worse: differential susceptibility to environmental influences. *Current Directions in Psychological Science*, *16*, 300–304.
- Belsky, D. W., & Israel, S. (2014). Integrating genetics and social science: genetic risk scores. *Biodemography and Social Biology*, *60*, 137–155.
- Belsky, J., & Pluess, M. (2009). Beyond diathesis stress: differential susceptibility to environmental influences. *Psychological Bulletin*, *135*, 885–908.
- Bolton, C., Barrowclough, C., & Calam, R. (2009). Parental criticism and adolescent depression: does adolescent self-evaluation act as a mediator? *Behavioural and Cognitive Psychotherapy*, *37*, 553–570.
- Brown, G. W., Birley, J. L., & Wing, J. K. (1972). Influence of family life on the course of schizophrenic disorders: a replication. *The British Journal of Psychiatry*, *121*, 241–258.
- Caspi, A., Sugden, K., Moffitt, T. E., Taylor, A., Craig, I. W., Harrington, H., & Poulton, R. (2003). Influence of life stress on depression: moderation by a polymorphism in the 5-HTT gene. *Science*, *301*, 386–389.
- Chang, C. C., Chow, C. C., Tellier, L. C. A. M., Vattikuti, S., Purcell, S. M., & Lee, J. J. (2015). Second-generation PLINK: rising to the challenge of larger and richer datasets. *GigaScience*, *4*, 7.
- Chen, J., Li, X., & McGue, M. (2013). The interacting effect of the BDNF Val66Met polymorphism and stressful life events on adolescent depression is not an artifact of gene-environment correlation: evidence from a longitudinal twin study. *Journal of Child Psychology and Psychiatry*, *54*, 1066–1073.
- Choi, S. W., Mak, T. S.-H., & O'Reilly, P. F. (2020). Tutorial: a guide to performing polygenic risk score analyses. *Nature Protocols*, *15*, 2759–2772.
- Cicchetti, D., & Rogosch, F. A. (2002). A developmental psychopathology perspective on adolescence. *Journal of Consulting and Clinical Psychology*, *70*, 6–20.
- Delaneau, O., Zagury, J.-F., & Marchini, J. (2013). Improved whole-chromosome phasing for disease and population genetic studies. *Nature Methods*, *10*, 5–6.
- Dick, D. M., Agrawal, A., Keller, M. C., Adkins, A., Aliev, F., Monroe, S., & Sher, K. J. (2015). Candidate gene-environment interaction research: reflections and recommendations. *Perspectives on Psychological Science*, *10*, 37–59.
- Dunn, E. C., Uddin, M., Subramanian, S. V., Smoller, J. W., Galea, S., & Koenen, K. C. (2011). Research review: Gene-environment interaction research in youth depression—a systematic review with recommendations for future research. *Journal of Child Psychology and Psychiatry*, *52*, 1223–1238.
- Eaton, W. W., Shao, H., Nestadt, G., Lee, B. H., Bienvenu, O. J., & Zandi, P. (2008). Population-based study of first onset and chronicity in major depressive disorder. *Archives of General Psychiatry*, *65*, 513–520.
- Garber, J., & Rao, U. (2014). Depression in children and adolescents. In M. Lewis & K. D. Rudolph (Eds.), *Handbook of Developmental Psychopathology* (pp. 489–520). New York, NY, US: Springer Science + Business Media.
- Hale, III, W. W., Raaijmakers, Q., Gerlisma, C., & Meeus, W. (2007). Psychometric properties of the Level of Expressed Emotion (LEE) questionnaire for adolescents from the general population. *Journal of Social Psychiatry and Psychiatric Epidemiology*, *42*, 215–220.
- Hammen, C., Brennan, P. A., Keenan-Miller, D., Hazel, N. A., & Najman, J. M. (2010). Chronic and acute stress, gender, and serotonin transporter gene-environment interactions predicting depression symptoms in youth. *Journal of Child Psychology and Psychiatry*, *51*, 180–187.
- Hankin, B. L., Fraley, R. C., Lahey, B. B., & Waldman, I. D. (2005). Is depression best viewed as a continuum or discrete category? A taxometric analysis of childhood and adolescent depression in a population-based sample. *Journal of Abnormal Psychology*, *114*, 96–110.
- Hankin, B. L., Young, J. F., Abela, J. R. Z., Smolen, A., Jenness, J. L., Gulley, L. D., & Oppenheimer, C. W. (2015). Depression from childhood into late adolescence: influence of gender, development, genetic susceptibility, and peer stress. *Journal of Abnormal Psychology*, *124*, 803–816.
- Harold, G. T., Rice, F., Hay, D. F., Boivin, J., Van Den Bree, M., & Thapar, A. (2011). Familial transmission of depression and antisocial behavior symptoms: disentangling the contribution of inherited and environmental factors and testing the mediating role of parenting. *Psychological Medicine*, *41*, 1175–1185.

- Hooley, J. M. (2007). Expressed emotion and relapse of psychopathology. *Annual Review of Clinical Psychology*, 3, 329–352.
- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Structural Equation Modeling*, 6, 1–55.
- Hyde, C. L., Nagle, M. W., Tian, C., Chen, X., Paciga, S. A., Wendland, J. R., & Winslow, A. R. (2016). Identification of 15 genetic loci associated with risk of major depression in individuals of European descent. *Nature Genetics*, 48, 1031–1036.
- Kessler, R. C., Angermeyer, M., Anthony, J. C., De Graaf, R., Demyttenaere, K., Gasquet, I., & Üstün, T. B. (2007). Lifetime prevalence and age-of-onset distributions of mental disorders in the World Health Organization's World Mental Health Survey Initiative. *World Psychiatry*, 6(3), 168–176.
- Kessler, R. C., Avenevoli, S., Costello, E. J., Georgiades, K., Green, J. G., Gruber, M. J., & Merikangas, K. R. (2012). Prevalence, persistence, and sociodemographic correlates of DSM-IV disorders in the National Comorbidity Survey Replication Adolescent Supplement. *Archives of General Psychiatry*, 69, 372–380.
- Kline, R. B. (2005). *Principles and practice of structural equation modeling*. 2nd edn. New York, NY: Guilford.
- Korn, J. M., Kuruvilla, F. G., McCarroll, S. A., Wysoker, A., Nemes, J., Cawley, S., & Altshuler, D. (2008). Integrated genotype calling and association analysis of SNPs, common copy number polymorphisms and rare CNVs. *Nature Genetics*, 40, 1253–1260.
- Lester, K. J., & Eley, T. C. (2013). Therapygenetics: using genetic markers to predict response to psychological treatment for mood and anxiety disorders. *Biology of Mood & Anxiety Disorders*, 3, 4.
- Li, Y., Willer, C. J., Ding, J., Scheet, P., & Abecasis, G. R. (2010). MaCH: using sequence and genotype data to estimate haplotypes and unobserved genotypes. *Genetic Epidemiology*, 34, 816–834.
- Liu, R. T. (2016). Taxometric evidence of a dimensional latent structure for depression in an epidemiological sample of children and adolescents. *Psychological Medicine*, 46, 1265–1275.
- McAdams, T. A., Neiderhiser, J. M., Rijdsdijk, F. V., Fruhling, V., Narusyte, J., Lichtenstein, P., & Eley, T. C. (2014). Accounting for genetic and environmental confounds in associations between parent and child characteristics: a systematic review of children-of-twins studies. *Psychological Bulletin*, 140, 1138–1173.
- McLaughlin, R. L., Schijven, D., van Rheenen, W., van Eijk, K. R., O'Brien, M., Kahn, R. S., & O'Donovan, M. C. (2017). Genetic correlation between amyotrophic lateral sclerosis and schizophrenia. *Nature Communications*, 8, 14774.
- Maske, U. E., Buttery, A. K., Beesdo-Baum, K., Riedel-Heller, S., Hapke, U., & Busch, M. A. (2016). Prevalence and correlates of DSM-IV-TR Major Depressive Disorder, self-reported diagnosed depression and current depressive symptoms among adults in Germany. *Journal of Affective Disorders*, 190, 167–177.
- McCarroll, S. A., Kuruvilla, F. G., Korn, J. M., Cawley, S., Nemes, J., Wysoker, A., & Altshuler, D. (2008). Integrated detection and population-genetic analysis of SNPs and copy number variation. *Nature Genetics*, 40, 1166–1174.
- McLeod, B. D., Weisz, J. R., & Wood, J. J. (2007). Examining the association between parenting and childhood depression: a meta-analysis. *Clinical Psychology Review*, 27, 986–1003.
- Merikangas, A. K., & Almasy, L. (2020). Using the tools of genetic epidemiology to understand sex differences in neuropsychiatric disorders. *Genes, Brain and Behavior*, 19, e12660.
- Monroe, S. M., & Simons, A. D. (1991). Diathesis-stress theories in the context of life stress research: implications for the depressive disorders. *Psychological Bulletin*, 110, 406–425.
- Muthén, L. K., & Muthén, B. O. (1998). *Mplus user's guide*. 8th edn. Los Angeles, CA: Muthén & Muthén. 2017.
- Naninck, E. F. G., Lucassen, P. J., & Bakker, J. (2011). Sex differences in adolescent depression: do sex hormones determine vulnerability? *Journal of Neuroendocrinology*, 3, 383–392.
- Nelemans, S. A., Hale, III, W. W., Branje, S. J. T., Hawk, S. T., & Meeus, W. H. J. (2014). Maternal criticism and adolescent depressive and generalized anxiety disorder symptoms: a 6-year longitudinal community study. *Journal of Abnormal Child Psychology*, 42, 755–766.
- Oldehinkel, A. J., & Bouma, E. M. C. (2011). Sensitivity to the depressogenic effect of stress and HPA-axis reactivity in adolescence: a review of gender differences. *Neuroscience and Biobehavioral Reviews*, 35, 1757–1770.
- Osman, A., Gutierrez, P. M., Bagge, C. L., Fang, Q., & Emmerich, A. (2010). Reynolds adolescent depression scale-second edition: a reliable and useful instrument. *Journal of Clinical Psychology*, 66, 1324–1345.
- Peris, T. S., & Miklowitz, D. J. (2015). Parental expressed emotion and youth psychopathology: new directions for an old construct. *Child Psychiatry and Human Development*, 46, 863–873.
- Pluess, M. (2015). Individual differences in environmental sensitivity. *Child Development Perspectives*, 9, 138–143.
- Pries, L.-K., Klingenberg, B., Menne-Lothmann, C., Decoster, J., van Winkel, R., Collip, D., & Guloksuz, S. (2020). Polygenic liability for schizophrenia and childhood adversity influences daily-life emotion dysregulation and psychosis proneness. *Acta Psychiatrica Scandinavica*, 141, 465–475.
- Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., & Sham, P. C. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. *American Journal of Human Genetics*, 81, 559–575.
- Purcell, S.M., Wray, N.R., Stone, J.L., Visscher, P.M., O'Donovan, M.C., Sullivan, P.F., & Sklar, P. (2009). Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature*, 460, 748–752. <https://doi.org/10.1038/nature08185>.
- Reynolds, W. M. (2000). *Reynolds Adolescent Depression Scale—2nd edition (RADS-2)*. Professional manual. Lutz, FL: Psychological Assessment Resources.
- Rohner, R. P., Khaleque, A., & Courmoyer, D. E. (2005). Parental acceptance-rejection: theory, methods, cross-cultural evidence, and implications. *Ethos*, 33, 299–334.
- Roisman, G. I., Newman, D. A., Fraley, R. C., Haltigan, J. D., Groh, A. M., & Haydon, K. C. (2012). Distinguishing differential susceptibility from diathesis-stress: recommendations for evaluating interaction effects. *Development and Psychopathology*, 24, 389–409.
- Rudolph, K. D. (2009). The interpersonal context of adolescent depression. In S. Nolen-Hoeksema & L. M. Hilt (Eds.), *Handbook of depression in adolescents* (pp. 377–418). New York, NY: Routledge.
- Scourfield, J., Rice, F., Thapar, A., Harold, G. T., Martin, N., & McGuffin, P. (2003). Depressive symptoms in children and adolescents: changing aetiological influences with development. *Journal of Child Psychology and Psychiatry*, 44, 968–976.
- Shih, J. H., Eberhart, N. K., Hammen, C. L., & Brennan, P. A. (2006). Differential exposure and reactivity to interpersonal stress predict sex differences in adolescent depression. *Journal of Clinical Child and Adolescent Psychology*, 35, 103–115.
- Sullivan, P. F., Neale, M. C., & Kendler, K. S. (2000). Genetic epidemiology of major depression: Review and meta-analysis. *American Journal of Psychiatry*, 157, 1552–1562.
- Tucker, G., Price, A. L., & Berger, B. (2014). Improving the power of GWAS and avoiding confounding from population stratification with PC-Select. *Genetics*, 197, 1045–1049.
- Van Assche, E., Moons, T., Cinar, O., Viechtbauer, W., Oldehinkel, A. J., Van Leeuwen, K., & Van Winkel, R. (2017). Gene-based interaction analysis shows GABAergic genes interacting with parenting in adolescent depressive symptoms. *Journal of Child Psychology and Psychiatry*, 58, 1301–1309.
- Van der Auwera, S., Peyrot, W., Milaneschi, Y., Hertel, J., Baune, B., Breen, G., & Grabe, H. (2018). Genome-wide gene-environment

interaction in depression: a systematic evaluation of candidate genes: The childhood trauma working-group of PGC-MDD. *American Journal of Medical Genetics, Part B: Neuropsychiatric Genetics*, 177, 40–49.

- Vaughn, C., & Leff, J. (1976). The measurement of expressed emotion in the families of psychiatric patients. *British Journal of Social and Clinical Psychology*, 15, 157–165.
- Widiger, T. A., & Samuel, D. B. (2005). Diagnostic categories or dimensions? A question for the Diagnostic and Statistical Manual of Mental Disorders—fifth edition. *Journal of Abnormal Psychology*, 114, 494–504.
- Wong, M. Y., Day, N. E., Luan, J. A., Chan, K. P., & Wareham, N. J. (2003). The detection of gene-environment interaction for continuous traits: should we deal with measurement error by bigger studies or better measurement? *International Journal of Epidemiology*, 32, 51–57.
- Wray, N. R., Lee, S. H., Mehta, D., Vinkhuyzen, A. A. E., Dudbridge, F., & Middeldorp, C. M. (2014). Research review: Polygenic methods and their application to psychiatric traits. *Journal of Child Psychology and Psychiatry*, 55, 1068–1087.
- Wray, N. R., Ripke, S., Mattheisen, M., Trzaskowski, M., Byrne, E. M., Abdellaoui, A., & Sullivan, P. F. (2018). Genome-wide association analyses identify 44 risk variants and refine the genetic architecture of major depression. *Nature Genetics*, 50, 668–681.

Stefanie A. Nelemans is an Assistant Professor at the Department of Youth and Family, Utrecht University, the Netherlands. Her major research interests include biological, psychological, and social factors associated with the development and maintenance of internalizing symptoms in adolescence and young adulthood.

Marco Boks is an Associate Professor at Department of Psychiatry, University Medical Center Utrecht, the Netherlands. His major research interests include genetic, epigenetic, and epidemiological studies of gene-environment interplay for psychiatric disorders such as PTSD, MDD, bipolar disorder, and schizophrenia.

Bochao Lin is a Postdoctoral Researcher at the Department of Translational Neuroscience, Brain Center Rudolf Magnus, the

Netherlands. Her major research interests include the etiology of psychiatric and neurological disorders considering genetic overlap between risk factors with diseases, side effects/drug response for antipsychotics, targeted next-generation sequencing, and single cell RNA sequencing.

Tineke Oldehinkel is Professor of Lifecourse Epidemiology of Common Mental Disorders at the University of Groningen, University Medical Center Groningen, the Netherlands. Her major research interests include the interplay of individual psychobiological vulnerability and resilience factors in the developmental course of depressive symptoms.

Pol van Lier is Professor at VU University Amsterdam, the Netherlands. His major research interests include social relationships in childhood and adolescence and their association with the development of psychopathology.

Susan Branje is Professor of Adolescent Development and chair of the Department of Youth and Family at Utrecht University, the Netherlands. Her major research interests include understanding the developmental changes in adolescents' interpersonal relationships and their associations with adolescent identity development and adjustment.

Wim Meeus is emeritus Professor at the Department of Youth and Family of Utrecht University, the Netherlands, and emeritus Professor at the Department of Developmental Psychology, Tilburg University, the Netherlands. His major research interests include adolescent development, identity and personality development, interpersonal relationships, and psychosocial problems in adolescence.