Single Nucleotide Polymorphism Heritability of Behavior Problems in Childhood: Genome-Wide Complex Trait Analysis

Irene Pappa, MD, MSc, Iryna O. Fedko, MSc, Viara R. Mileva-Seitz, PhD, Jouke-Jan Hottenga, PhD, Marian J. Bakermans-Kranenburg, PhD, Meike Bartels, PhD, Catharina E.M. van Beijsterveldt, PhD, Vincent W.V. Jaddoe, MD, PhD, Christel M. Middeldorp, MD, PhD, Ralph C.A. Rippe, PhD, Fernando Rivadeneira, MD, PhD, Henning Tiemeier, MD, PhD, Frank C. Verhulst, MD, PhD, Marinus H. van IJzendoorn, PhD, Dorret I. Boomsma, PhD

Objective: Genetic factors contribute to individual differences in behavior problems. In children, genome-wide association studies (GWAS) have yielded the first suggestive results when aiming to identify genetic variants that explain heritability, but the proportion of genetic variance that can be attributed to common single nucleotide polymorphisms (SNPs) remains to be determined, as only a few studies have estimated SNP heritability, with diverging results.

Method: Genomic-relationship-matrix restricted maximum likelihood (GREML) as implemented in the software Genome-Wide Complex Trait Analysis (GCTA) was used to estimate SNP heritability (SNP h^2) for multiple phenotypes within 4 broad domains of children's behavioral problems (attention-deficit/hyperactivity symptoms, internalizing, externalizing, and pervasive developmental problems) and cognitive function. We combined phenotype and genotype data from 2 independent, population-based Dutch cohorts, yielding a total number of 1,495 to 3,175 of 3-, 7-, and 9-year-old children.

Results: Significant SNP heritability estimates were found for attention-deficit/hyperactivity symptoms (SNP $h^2 = 0.37–0.71$), externalizing problems (SNP $h^2 = 0.44$), and total problems (SNP $h^2 = 0.18$), rated by mother or teacher. Sensitivity analyses with exclusion of extreme cases and quantile normalization of the phenotype data decreased SNP h^2 as expected under genetic inheritance, but they remained statistically significant for most phenotypes.

Conclusion: We provide evidence of the influence of common SNPs on child behavior problems in an ethnically homogenous sample. These results support the continuation of large GWAS collaborative efforts to unravel the genetic basis of complex child behaviors.

Key Words: genome-wide complex trait analysis (GCTA), heritability, children, behavior problems

J Am Acad Child Adolesc Psychiatry 2015;54(9):737–744.

Complex behaviors are shaped by both genetic and environmental influences,[1,2] and numerous twin, family, and adoption studies have estimated significant contributions of genetic factors to individual differences in behavioral and psychiatric traits.[3-5] In addition, longitudinal population-based studies provide evidence of the genetic stability of common behavioral problems (e.g., anxiety and depression symptoms,[6] attention problems[7]) across the lifespan, with higher heritability estimates in childhood (e.g., for attention problems, heritability estimates decreased from 0.70 in childhood to 0.40 in adulthood[7]).

In adult samples, genome-wide association studies (GWAS) identified genes and pathways related to complex traits.[8,9] This approach has also yielded positive findings in studies of important traits in children (e.g., birth weight[10] and length[11]). For childhood psychiatric traits and problem behaviors, successes have been limited,[12-15] which can probably be ascribed to the very modest sample sizes in these studies.[16] The relatively small or absent genetic associations with complex traits of interest in GWAS[12-15] may seem in contrast to the large heritability estimates from twin and family studies but are indeed in line with recent evidence that the small effect sizes of individual SNPs may be responsible for the nonreplicability of these associations.[17] To assess whether GWAS of child behavior problems can be expected to yield important findings regarding biological pathways, we address the question of what part of the heritability of childhood behavior problems is captured by common (minor allele frequency >1%) single nucleotide polymorphisms (SNPs) included in standard genotyping arrays.

The genetic variance explained by genome-wide SNPs[18] can be estimated by using the genetic similarity among unrelated individuals as a predictor of their phenotypic resemblance. When individual-level genotype data are available, these can be used to obtain a measure of genetic similarity between all possible pairs of (unrelated)

This article is discussed in an editorial by Drs. Philip Asherson and Paul F. O'Reilly on page 702.

Supplemental material cited in this article is available online.

individuals in the study. In a second step, this genetic relatedness matrix (GRM) is used to predict the phenotype similarity between individuals, just as the different similarity of monozygotic (MZ) and dizygotic (DZ) twin pairs predicts their different phenotype resemblances. This approach has been implemented in the software package Genome-Wide Complex Trait Analysis (GCTA).[18] The heritability estimates from GCTA (SNP h^2) are commonly considered an indicator of the upper limit of the variance that can be explained by current GWAS efforts. Power estimations have indicated that for quantitative traits, a sample size of 3,000 individuals is required to detect an SNP h^2 of 0.30 with 80% power.[19] Thus, large sample sizes are required to reliably estimate the SNP heritability of complex behavioral traits, which can imply the need to pool data from multiple studies.

To date, few SNP heritability estimates are available for behavioral problems in childhood. Some studies indicate substantial additive genetic heritability of normative differences in children's social communication difficulties[20] and in clinical cases of attention-deficit/hyperactivity disorder (ADHD)[21,22] and childhood-onset obsessive-compulsive disorder (OCD).[23] However, other studies indicate modest, statistically nonsignificant SNP heritability estimates for children's internalizing problems,[12] anxiety,[24] and callous-unemotional (CU) traits[13] in population-based samples. A study from the Twins Early Development Study (TEDS) indicated no significant SNP heritability for parent-, teacher-, and self-reported behavioral problems (i.e., attention problems, internalizing, and externalizing problems) in contrast to cognitive and anthropomorphic traits in a population-based sample (n = 2,500) of 12-year-old children.[25]

Here we focus on 4 domains of children's behavioral problems: attention deficit problems, externalizing, internalizing, and pervasive developmental problems. Genetic influences on nonverbal cognitive abilities were also estimated. To obtain sufficient power, we combined genotype and phenotype data from 2 independent, population-based Dutch cohorts: the Generation R Study (GEN-R) and the Netherlands Twin Register (NTR). Genotyped SNP data from both studies were used to construct a GRM.[26] For both studies, behavior problems of a total number of 1,495 to 3,175 of 3-, 7-, and 9-year-old children were rated by mothers and/or teachers. We estimated the SNP heritability in each of these traits, and we compared our findings to the SNP heritability estimates previously reported.

METHOD

Participants

This study included data from children from 2 population-based Dutch cohorts, GEN-R and NTR. GEN-R is a prospective cohort based in Rotterdam. The characteristics of the study have been previously described in detail.[27] NTR is a nationwide longitudinal sample of twins and their family members followed from birth onward after voluntary registration.[28] In both studies, parents gave informed consent for participation and also to approach teachers of the children. Study protocols were approved by the local ethics committees.

Measures

All phenotypes analyzed in this study have been described in detail in previous publications of GEN-R and NTR, and twin-based heritabilities in the Dutch population were reported for these traits (see Table S1, available online).

Conners' Parent Rating Scale. ADHD symptoms and related comorbid symptoms were assessed using the Conners' Parent Rating Scale (CPRS-R)[26] completed by the mothers. Four scales of the CPRS-R were used: ADHD Combined; ADHD Inattentive; ADHD Hyperactive-Impulsive; and Oppositional Defiant Disorder (ODD) scale.

Child Behavior Checklist. We assessed child behavior problems using the well-validated Child Behavior Checklist (CBCL),[27] completed by the mother. Internalizing, externalizing, and total problems were assessed using the appropriate CBCL syndrome scales. For the CBCL Internalizing, Externalizing, and Total Problems scores, the GEN-R study used the CBCL for ages 1.5 to 5 years,[28] and NTR used the CBCL for ages 6 to 18 years.[29] In the CBCL for ages 1.5 to 5 years, the Internalizing scale consists of 4 scales (Emotionally Reactive, Anxious/Depressed, Somatic Complaints, and Withdrawn), and the Externalizing scale consists of 2 scales (Attention Problems and Aggressive Behavior). In the CBCL for ages 6 to 18 years, the Internalizing scale consists of 3 scales (Anxious/Depressed, Withdrawn/Depressed, and Somatic Complaints), and the Externalizing scale consists of 2 scales (Rule-Breaking Behavior and Aggressive Behavior). The Total Problems score is computed by summing the ratings of all problem items included in the CBCL. To avoid phenotypic heterogeneity in the combined dataset due to differences in the items between the 2 CBCL versions, we selected only overlapping items to compute the scores (see Table S2, available online).

We assessed pervasive developmental problems using the Pervasive Developmental Disorder (PDD) subscale of the CBCL for children 1.5 to 5 years.[28] The PDD subscale has been shown to be a valid screening tool for autism spectrum disorders (ASD).[30]

Teacher's Rating Form. The Teacher's Rating Form (TRF) for ages 6 to 18 years[29] was used to assess attention problems (Attention Problems scale) and behavioral problems (Externalizing scale), rated by the teacher. We used the teachers' ratings of externalizing and not internalizing problems, since it has been previously shown that they can better identify children with externalizing rather than internalizing problems.[31] The teacher reports were also selected to assess behavior in a different environment, and to avoid informant effects, which could bias estimates of genetics contribution to common child behavior problems.[32,33]

Nonverbal Cognitive Abilities. Nonverbal cognitive abilities were assessed with the Snijder–Oomen Nonverbal Intelligence Test[34] (SON-R 2.5–7 years) in the GEN-R study, and the nonverbal subtest of the Revised Amsterdam Children Intelligence Test[35] (RAKIT) in the NTR. Both measurements are well validated and correlate substantially with the Wechsler Preschool and Primary Scale of Intelligence–Revised (WPPSI-R)[36] and the Wechsler Intelligence Scale for Children (WISC).[37] The nonverbal cognition scores in both samples were transformed to mean = 100 and SD = 15.

Genotyping and Imputation

A total of 3,102 children from the GEN-R study and 2,826 children from NTR, all of white ethnicity, were genotyped on Illumina (660W, 610K) and Affymetrix 6.0 platforms, respectively. Because the number of overlapping SNPs between platforms was small ($n \sim 120K$), both cohorts were cross-platform imputed using MaCH-Admix imputation software[38] as described in Fedko *et al.*[26] Cross-platform imputation supplies all participants from both cohorts with genetic information from all SNPs genotyped on both

platforms. To avoid population stratification between samples due to genotyping platform, the Genome of the Netherlands reference set[39] was used to phase and subsequently impute missing genotypes into both cohorts. The final dataset consisted of 5,928 individuals where each individual had information for 989,757 SNPs expressed in dosage scores. Postimputation quality control (QC) was performed on imputed datasets to check and control for possible residual imputation stratification due to genotyping platform or true genetic differences between cohorts. The overall imputation quality measure (R^2) was high (mean = 0.97, median = 0.99). Case-control analysis of the imputed sample, where GEN-R children were assigned as participants and NTR children as controls, showed 4,340 SNPs that were significantly different in frequency ($p < 10^{-5}$). These SNPs were excluded from further analysis.

Genome-Wide Complex Trait Analysis (GCTA)

We built a GRM based on cross-platform imputed data using GCTA version 1.20.[40] Data for the GRM was filtered based on the following 2 criteria: $R^2 > 0.8$ to allow SNPs with high imputation quality; and minor allele frequency > 0.01 to exclude SNPs with low minor allele frequency. We performed principal components analysis (PCA) on the resulting GRM to check for possible residual stratification due to genotyping platform. We used the GREML (Genomic-relatedness-matrix restricted maximum likelihood) method to estimate SNP heritability in distantly related individuals from all genotyped and imputed SNPs in the dataset. The convention excludes those individuals whose genetic relatedness exceeds the 0.025 threshold in GRM, which corresponds to third- or fourth-degree cousins. We applied such a cut-off while performing GREML analysis, and 1 of each pair of closely related individuals was excluded from analysis, resulting in a number range from 1,495 to 3,175 depending on phenotype (see Table S4, available online). For all phenotypes, we included age and sex as covariates. We also adjusted for the cohort of origin (GEN-R or NTR) to control for residual imputation stratification due to genotyping platform, true genetic differences, and possible phenotype differences.

Statistical Analyses

In both GEN-R and NTR, nonresponse analysis indicated no differences in the baseline characteristics of children whose assessment of child behavior problems was not completed at ages 7 and 9 years. As these are longitudinal studies, families are permitted to miss participation for a particular survey, allowing them to participate again in later surveys.[41] To explore the effect of extreme cases, often found in ratings of children's behavior problems, we winsorized phenotypes used in this study when it was required. If the corresponding absolute z score was more than 3.29 for a phenotype, we replaced the raw score with the less extreme value, that is, with the next highest score plus 1 unit.[42] In addition, we checked for possible population stratification in combined dataset, adjusting for 10 principal components (PCs) in analysis of each scale.

Sensitivity analyses were also performed to explore the influence of exclusion of extreme cases, skewness on SNP heritability estimates, and the influence of study of origin. First, we estimated SNP heritability by excluding extreme scores below or above 3 standard deviations from the mean. If such cases represent extremes due to, for example, measurement errors, we expect SNP heritability to increase after exclusion of these cases. If, however, they represent genuine outliers, we expect them also to be outliers for heritable traits, and consequently the SNP heritability will decrease after exclusion of these cases. Second, we transformed the data to the quantile normalized scale, using the Van der Waerden transformation. This transformation reduces the extreme influence that outliers could have by ranking them as low or high within a normal distribution,[43] although it results in some loss of phenotypic information. In addition, we performed GCTA separately on the 2 participating studies (i.e., GEN-R and NTR) to explore possible effects of the specific study. All transformations were conducted in SPSS 21.0 software.[44]

RESULTS

Genotypic and Phenotypic Sample Characteristics

The sample characteristics of the children participating in each study and in the combined dataset, before and after exclusion of related individuals, are presented in Tables S3 and S4, available online. The distribution of age, sex, and behavior problems did not significantly differ between the 2 studies.

Estimates of SNP Heritability

Table 1 summarizes the SNP heritability estimates using the combined GRM, adjusting for age, sex, and sample of origin. For the mother ratings of child problem behavior, estimates were substantial and statistically significant for the ADHD Combined scale (SNP $h^2 = 0.40$, SE = 0.14, $p = .001$), as well as for the ADHD Inattentive scale (SNP $h^2 = 0.37$, SE = 0.14, $p = .003$) and the Hyperactive-Impulsive scale (SNP $h^2 = 0.45$, SE = 0.14, $p = .0006$), measured by the CPRS. We also found significant SNP heritability estimates for the CBCL Total Problems score (SNP $h^2 = 0.18$, SE = 0.10, $p = .03$). For the teacher ratings, we obtained significant SNP heritability estimates for both the Attention Problems scale (SNP $h^2 = 0.71$, SE = 0.22, $p = .0006$) and the Externalizing scale (SNP $h^2 = 0.44$, SE = 0.22, $p = .03$). No significant estimates were found for the CPRS ODD scale, the CBCL PDD subscale, the CBCL Internalizing and Externalizing scales, and the nonverbal cognition. When 10 PCs were used as covariates, we found nonsignificant difference in SNP heritability estimates (1%–3% drop or 1%–2% increase, data not shown) for all scales. The level of significance remained the same, except for CBCL Total Problems score ($p = .07$ and $p = .03$ with and without PCs adjustment, accordingly).

Sensitivity Analyses

To examine the influence of extreme cases on SNP heritability estimates and to enable comparison with the TEDS results[25] that were based on deletion of extreme cases, we also performed GCTA analyses excluding individuals above or below 3 standard deviations from the mean. Overall, the exclusion of extreme cases decreased the SNP heritability estimates by almost half in most of the complex problem behaviors, suggesting that these children are genuine outliers and that their extreme phenotype values do not represent measurement errors or other artifacts. Even after removal of outliers, SNP heritability estimates for all scales of CPRS (ADHD Combined, ADHD Inattentive, ADHD Hyperactive-Impulsive, and ODD scale) were still substantial. The CBCL Total Problems score and the teacher-reported Attention Problems and Externalizing

TABLE 1 Single Nucleotide Polymorphisms (SNP) Heritability Estimates in Child Behavior Problems

	Age (SD)	SNP h^2	SE	95% CI[a]	n	*p*
Parent Ratings						
CPRS ADHD Combined scale	8.34 (0.7)	0.40	0.14	(0.13, 0.67)	2,262	<.01**
CPRS ADHD Inattentive scale	8.34 (0.7)	0.37	0.14	(0.10, 0.64)	2,264	<.01**
CPRS Hyperactive-Impulsive scale	8.34 (0.7)	0.45	0.14	(0.18, 0.72)	2,260	<.001***
CPRS ODD scale	8.34 (0.7)	0.20	0.14	(0.00, 0.47)	2,262	.07
CBCL Internalizing scale	6.57 (0.83)	0.12	0.10	(0.00, 0.32)	3,175	.11
CBCL Externalizing scale	6.57 (0.83)	0.12	0.10	(0.00, 0.32)	3,174	.13
CBCL Total Problems score	6.57 (0.83)	0.18	0.10	(0.00, 0.38)	3,175	<.05*
CBCL PDD subscale	3.15 (0.23)	0.16	0.11	(0.00, 0.33)	3,015	.07
Teacher Ratings						
TRF Attention Problems scale	6.82 (2.35)	0.71	0.22	(0.28, 1.00)	1,495	<.001***
TRF Externalizing scale	6.82 (2.35)	0.44	0.22	(0.01, 0.87)	1,495	<.05*
Observational Ratings						
Nonverbal cognition	6.14 (0.42)	0.11	0.16	(0.00, 0.42)	1,974	.23

Note: All analyses were performed with the combined genetic relatedness matrix and were adjusted for age, sex, and sample of origin (the Generation R Study or the Netherlands Twin Register) on winsorized scores. ADHD = attention-deficit/hyperactivity disorder; CBCL = Child Behavior Checklist; CPRS = Conners' Parent Rating Scale; ODD = oppositional defiant disorder; PDD = pervasive developmental disorder; SE = standard error; TRF = Teacher's Rating Form.

[a]*SNP heritability estimates are limited to (0.00-1.00).*

*p < .05; **p < .01; ***p < .001.

Problems also remained significant after exclusion of extreme cases. The results are summarized in Table 2.

In addition, similar to the TEDS study,[25] we performed GCTA analyses on the quantile normalized scales, using the Van der Waerden transformation to examine the potential influence of skewness on SNP heritability estimates. The SNP heritability with the transformed scales was similar to those with untransformed scales, with substantial genetic effects contributing to the ADHD Combined scale (SNP $h^2 = 0.30$, SE = 0.14, $p = .01$), ADHD Inattentive scale (SNP $h^2 = 0.30$, SE = 0.14, $p = .01$), and Hyperactivity-Impulsive scale (SNP $h^2 = 0.37$, SE = 0.14, $p = .004$) rated by the mother using the CPRS. Also, the CBCL PDD subscale (SNP $h^2 = 0.18$, SE = 0.11, $p = .05$) and the teacher-reported Attention Problems scale (SNP $h^2 = 0.64$, SE = 0.22, $p = .002$) and Externalizing scale (SNP $h^2 = 0.60$, SE = 0.22, $p = .004$) yielded significant SNP heritability estimates. The results are summarized in Table S5, available online.

Finally, we also provide SNP heritability estimates of the 2 samples independently. As expected, the smaller NTR sample shows estimates with larger SE values. Although variable, the SNP heritability estimates of the 2 samples did not differ significantly from each other. The results are summarized in Table S6, available online.

DISCUSSION

The aim of this study was to provide estimates of SNP heritability of normative differences in attention deficit problems (measured at age 7 and 9 years), externalizing and internalizing problems (measured at 7 years), pervasive developmental problems (measured at 3 years), and nonverbal cognitive function (measured at 7 years) in population-based samples. Our study provides evidence of significant SNP heritability for attention-deficit/hyperactivity problems, externalizing, and total problems, rated by mother or teacher. We identified nonsignificant SNP heritability estimates for pervasive developmental and internalizing problems. These results are parallel to twin heritabilities previously reported on the same phenotypes, that is, higher twin heritabilities were associated with higher SNP heritabilities. Sensitivity analyses showed that SNP heritability estimates decreased but remained significant for most phenotypes after exclusion of the extreme cases.

Previous studies on the heritability captured by common SNPs have yielded significant SNP heritability estimates for normative differences in autistic-like traits[20,45] and in clinical cases of childhood-onset OCD[23] and ADHD.[21,22] Quantifiable although nonsignificant SNP heritability has also been reported for internalizing problems in population-based samples of preschoolers.[12] However, a recent TEDS study by Trzaskowski *et al.*[25] indicated no additive genetic effects for common child behavior problems. This discrepancy may be due to several factors. First, there are methodological differences between the 2 studies. In the present study, we removed ethnic outliers instead of correcting for them by principal components. Moreover, we estimated heritability with and without extreme cases, showing that in some cases, treatment of outliers results in substantially different findings. Extreme cases might be biologically significant extremes, or they might constitute statistical outliers. Our results indicate that extremes were more likely to be genetic extremes rather than statistical outliers, and they suggest that in the TEDS study,[25] the exclusion of extreme cases may have resulted in an underestimation of SNP heritability.[46] Winsorizing the extreme cases instead of excluding them may address the problem of extremely skewed distributions while still retaining information for all participants. It should

TABLE 2 Impact of Extreme Cases on Single Nucleotide Polymorphisms (SNP) Heritability Estimates in Child Behavior Problems

	Age (SD)	SNP h^2	SE	95% CI[a]	n	*p*
Parent Ratings						
CPRS ADHD Combined scale	8.34 (0.7)	0.22	0.14	(0.00, 0.49)	2,240	.05*
CPRS ADHD Inattentive scale	8.34 (0.7)	0.24	0.14	(0.00, 0.51)	2,229	<.05*
CPRS Hyperactive-Impulsive scale	8.34 (0.7)	0.33	0.15	(0.04, 0.62)	2,231	.01*
CPRS ODD scale	8.34 (0.7)	0.28	0.14	(0.01, 0.55)	2,246	<.05*
CBCL Internalizing scale	6.57 (0.83)	0.04	0.10	(0.00, 0.24)	3,139	.36
CBCL Externalizing scale	6.57 (0.83)	0.06	0.10	(0.00, 0.26)	3,136	.28
CBCL Total Problems score	6.57 (0.83)	0.16	0.10	(0.00, 0.36)	3,143	.05*
CBCL PDD subscale	3.15 (0.23)	0.14	0.11	(0.00, 0.36)	2,999	.10
Teacher Ratings						
TRF Attention Problems scale	6.82 (2.35)	0.49	0.22	(0.06, 0.92)	1,470	.01*
TRF Externalizing scale	6.82 (2.35)	0.46	0.23	(0.01, 0.91)	1,463	<.05*
Observational Ratings						
Nonverbal cognition	6.14 (0.42)	0.11	0.16	(0.00, 0.42)	1,968	.23

Note: In all analyses, statistical outliers (mean ±3 SD) were excluded. All analyses were performed with the combined genetic relatedness matrix and were adjusted for age, sex, and sample of origin (the Generation R Study or the Netherlands Twin Register). ADHD = attention-deficit/hyperactivity disorder; CBCL = Child Behavior Checklist; CPRS = Conners' Parent Rating Scale; ODD = oppositional defiant disorder; PDD = pervasive developmental disorder; SE = standard error; TRF = Teacher's Rating Form.

[a] *SNP heritability estimates are limited to (0.00–1.00).*

*p < .05.

be noted, however, that even after exclusion of the extreme cases, we found significant additive genetic heritability for ADHD-related symptoms and children's behavior problems.

Second, the 2 studies involve different samples. The TEDS sample[25] involved 12-year-old children, whereas in our sample, we analyzed data on behavior problems at 3, 7, and 9 years. Estimations of genetic effects may differ developmentally, although the direction depends on the phenotype of interest. For example, SNP heritability of autistic-like traits was shown to be low (SNP h^2 = 0.24, SE = 0.07, n = 5,204) but strongest in early childhood,[20] whereas it increased from age 7 to 12 years in the case of general cognitive ability.[47] Previous twin studies have also indicated an increase in the heritability for general cognitive ability,[48] as well as for nonverbal IQ[49] from childhood to adulthood. In line with the low heritability estimates in 5- to 6-year-old twins for nonverbal IQ, we found no significant SNP heritability of nonverbal cognitive ability at 6 years in a subsample of 1,974 unrelated individuals. Similarly, our study indicated nonsignificant SNP heritability of pervasive developmental problems in 3-year-old children (n = 3,015). Given the low overall heritability, larger samples may be needed to estimate modest SNP heritability of nonverbal cognition and pervasive developmental problems in early childhood. The perception of genetic heritability as time/age dependent[50] could explain discrepancies between samples and between measurements at different time points (e.g., CBCL measures at 7 years and CPRS measures at 9 years), and suggests that SNP heritability estimates cannot be easily generalized across age.

Third, SNP heritability, as an estimation of the fraction of phenotypic variation explained by common SNPs, is dependent on sample characteristics.[51] Thus, as a population property, SNP heritability estimates can differ between samples, both because the GRM differs between populations and because environmental factors are different. Environmental influences may play a more important role in a sample derived from multiple, culturally diverse sites in the United Kingdom, whereas genetic effects would be more prominent in the geographically restricted and rather homogeneous Dutch society in terms of socioeconomic conditions. Parental reports of child problem behaviors might be determined partly by subjective criteria for what parents consider to be problem behavior, and these criteria may be dependent on cultural norms or socioeconomic circumstances such as crowding.[52]

In this study, we found nonsignificant SNP heritability for parent-reported internalizing problems in 7-year-old children. One reason for this finding could be the difficulty in assessing internalizing symptoms in early childhood. Internalizing symptoms are often not overtly expressed in young children and thus not easily observed by the parents.[53] Another reason could be that since the prevalence of internalizing symptoms typically increases in middle-to-late adolescence,[54,55] we are not yet able to identify all children who will develop internalizing symptoms later in life. The particularly high heterogeneity and the distinctive genetic architecture of internalizing problems have also been addressed in previous work.[56]

A limitation of this study is the sample size. GCTA power calculations indicate that even with large sample sizes, the SEs of the SNP heritability estimates are large.[19] Thus, even larger samples are needed to estimate modest additive genetic effects. However, the sample size of the current study is for most phenotypes comparable to the study of Trzaskowski *et al.*,[25] indicating that sample size is not exclusively responsible for the discrepancies between studies. A parameter inherent to most behavior problems

research is the skewed distribution of the phenotypes. Nevertheless, sensitivity analyses with transformed distributions and winsorized extreme cases did not reduce the significant SNP heritability estimates to nonsignificance. This study is based on data from 2 longitudinal studies (GEN-R and NTR). Systematic attrition is a limitation inherent to longitudinal studies,[57] potentially leading to selective dropout of high-risk individuals and thus to underestimation of the heritability of common behavior problems in children. However, previous research has shown that psychopathology of the participants has a small to moderate effect on attrition rates[58,59] and that estimations from longitudinal studies are robust and generalizable.[60,61] Finally, the results of this study are derived from population-based samples of children. Although it has been shown that additive effects of hundreds of SNPs are responsible for observed normal variation in most quantitative traits,[62] it is possible that the genetic architecture of children diagnosed with severe behavior problems differs from that of children in population-based samples (e.g., increased role of rare variants and de novo mutations).

In summary, this study provides molecular genetic evidence of additive genetic influences on specific child behavior problems in an ethnically and socioeconomically homogeneous sample. SNP heritability for other common behavior problems in children, or for the *p* factor as proposed by Caspi *et al.*,[63] remains to be estimated. SNP heritability estimates may be influenced by diversity in socioeconomic environment, developmental stage, and study design, arguing for approaches that model gene-by-environment interactions, developmental information, and possibly data from population-based and clinical samples in GCTA research. Our results provide support for and encourage the continuation of GWAS efforts by genetics consortia focusing on complex behavioral traits in search of elusive heritability. &

Accepted June 16, 2015.

Drs. Pappa and van IJzendoorn are with the School of Pedagogical and Educational Sciences, Erasmus University Rotterdam, Rotterdam, The Netherlands and the Generation R Study Group, Erasmus Medical Center. Dr. van IJzendoorn is also with the Centre for Child and Family Studies, Leiden University, Leiden, The Netherlands. Ms. Fedko and Drs. Hottenga, Bartels, van Beijsterveldt, Middeldorp, and Boomsma are with VU University Amsterdam. Drs. Hottenga, Bartels, and Boomsma are also with the EMGO+ Institute for Health and Care Research, VU University Medical Center, Amsterdam. Drs. Bartels, Middeldorp, and Boomsma are also with the Neuroscience Campus Amsterdam. Drs. Mileva-Seitz, Jaddoe, Tiemeier, and Verhulst are with the Generation R Study Group, Erasmus Medical Center. Drs. Mileva-Seitz, Bakermans-Kranenburg, and Rippe are with the Centre for Child and Family Studies. Drs. Jaddoe, Rivadeneira, and Tiemeier are with Erasmus Medical Center. Drs. Tiemeier and Verhulst are with Erasmus Medical Center-Sophia Children's Hospital, Rotterdam.

The Generation R Study is conducted by the Erasmus Medical Center in close collaboration with the Erasmus University Rotterdam, School of Law and Faculty of Social Sciences, the Municipal Health Service Rotterdam area, Rotterdam, the Rotterdam Homecare Foundation, Rotterdam, and the Stichting Trombosedienst en Artsenlaboratorium Rijnmond (STAR), Rotterdam. The authors gratefully acknowledge the contribution of general practitioners, hospitals, midwives, and pharmacies in Rotterdam. The first phase of the Generation R Study is made possible by financial support from: Erasmus Medical Centre, Rotterdam, Erasmus University Rotterdam, and the Netherlands Organization for Health Research and Development (ZonMw). H. Tiemeier, MD, PhD, was supported by NOW-ZonMW VIDI (grant no. 017.106.370). M.H. van IJzendoorn, PhD, and M.J. Bakermans-Kranenburg, PhD, were supported by the Dutch Ministry of Education, Culture, and Science and the Netherlands Organization for Scientific Research (Gravitation program, SPINOZA, VICI). Viara Mileva-Seitz, PhD, was supported by NWO RUBICON prize no. 446-11-023. M.J. Bakermans-Kranenburg, PhD, H. Tiemeier, MD, PhD, F.C. Verhulst, MD, PhD, and M.H. van IJzendoorn, PhD, are also members of the Consortium on Individual Development, which is funded through the Gravitation program of the Dutch Ministry of Education, Culture, and Science and the Netherlands Organization for Scientific Research (NWO grant no. 024.001.003).

The Netherlands Twin Register thanks all twins, their parents, and teachers for their participation and acknowledges support from the American Recovery and Reinvestment Act (ARRA) project RC22MH08995; the Netherlands Organization for Scientific Research "Spinozapremie" (NWO/SPI 56-464-14192); Twin-family database for behavior genetics and genomics studies (NWO 480-04-004); Genetic influences on stability and change in psychopathology from childhood to young adulthood (ZonMW 912-10-020); NWO Center for Medical Systems Biology (CMSB); the Netherlands Bioinformatics Centre (NBIC) Bioassist (NWO-NBIC/BioAssist/RK/2008.024); the Avera Institute for Human Genetics, Sioux Falls, South Dakota (USA), Biobanking and Biomolecular Resources Research Infrastructure Rainbow project 1 and 2 (NWO 184.021.007). The authors are grateful to NTR data collection and data management team. The authors thank all the Generation R study (GEN-R) and Netherlands Twin Registry (NTR) participants, who contributed to this study.

Disclosure: Dr. Verhulst has received remuneration as contributing author of the Achenbach System of Empirically Based Assessment (ASEBA). Drs. Pappa, Mileva-Seitz, Hottenga, Bakermans-Kranenburg, Bartels, van Beijsterveldt, Jaddoe, Middeldorp, Rippe, Rivadeneira, Tiemeier, van IJzendoorn, Boomsma, and Ms. Fedko report no biomedical financial interests or potential conflicts of interest.

Correspondence to Marinus H. van IJzendoorn, PhD, Centre for Child and Family Studies, Leiden University, PO Box 9555, 2300 RB Leiden, the Netherlands; e-mail: vanijzen@fsw.leidenuniv.nl

0890-8567/$36.00/©2015 American Academy of Child and Adolescent Psychiatry

http://dx.doi.org/10.1016/j.jaac.2015.06.004

REFERENCES

1. Bouchard TJ Jr, McGue M. Genetic and environmental influences on human psychological differences. J Neurobiol. 2003;54:4-45.
2. McGue M, Bouchard TJ Jr. Genetic and environmental influences on human behavioral differences. Annu Rev Neurosci. 1998;21:1-24.
3. Hicks BM, Krueger RF, Iacono WG, McGue M, Patrick CJ. Family transmission and heritability of externalizing disorders—a twin-family study. Arch Gen Psychiatry. 2004;61:922-928.
4. Bartels A, van den Oord EJCG, Hudziak JJ, Rietveld AJH, van Beijsterveldt CEM, Boomsma DI. Genetic and environmental mechanisms underlying stability and change in problem behaviors at ages 3, 7, 10, and 12. Dev Psychol. 2004;40:852-867.
5. Bezdjian S, Baker LA, Tuvblad C. Genetic and environmental influences on impulsivity: a meta-analysis of twin, family and adoption studies. Clin Psychol Rev. 2011;31:1209-1223.
6. Nivard MG, Dolan CV, Kendler KS, *et al*. Stability in symptoms of anxiety and depression as a function of genotype and environment: a longitudinal twin study from ages 3 to 63 years. Psychol Med. 2015;45:1039-1049.
7. Kan KJ, Dolan CV, Nivard MG, *et al*. Genetic and environmental stability in attention problems across the lifespan: evidence from the Netherlands Twin Register. J Am Acad Child Psychiatry. 2013;52:12-25.
8. Visscher PM, Brown MA, McCarthy MI, Yang J. Five years of GWAS discovery. Am J Hum Genet. 2012;90:7-24.
9. Rietveld CA, Medland SE, Derringer J, *et al*. GWAS of 126,559 individuals identifies genetic variants associated with educational attainment. Science. 2013;340:1467-1471.
10. Horikoshi M, Yaghootkar H, Mook-Kanamori DO, *et al*. New loci associated with birth weight identify genetic links between intrauterine

growth and adult height and metabolism. Nature Genet. 2013;45:76-U115.
11. van der Valk RJ, Kreiner-Moller E, Kooijman MN, *et al*. A novel common variant in DCST2 is associated with length in early life and height in adulthood. Hum Mol Genet. 2014;24:1155-1168.
12. Benke KS, Nivard MG, Velders FP, *et al*. A genome-wide association meta-analysis of preschool internalizing problems. J Am Acad Child Adolesc Psychiatry. 2014;53:667-676.
13. Viding E, Price TS, Jaffee SR, *et al*. Genetics of callous-unemotional behavior in children. PLoS One. 2013;8:e65789.
14. Lasky-Su J, Neale BM, Franke B, *et al*. Genome-wide association scan of quantitative traits for attention deficit hyperactivity disorder identifies novel associations and confirms candidate gene associations. Am J Med Genet B Neuropsychiatr Genet. 2008;147B:1345-1354.
15. Mick E, McGough J, Loo S, *et al*. Genome-wide association study of the child behavior checklist dysregulation profile. J Am Acad Child Adolesc Psychiatry. 2011;50:807-817.
16. Wray NR, Lee SH, Mehta D, Vinkhuyzen AAE, Dudbridge F, Middeldorp CM. Research review: polygenic methods and their application to psychiatric traits. J Child Psychol Psychiatry. 2014;55:1068-1087.
17. Rietveld CA, Conley D, Eriksson N, *et al*. Replicability and robustness of genome-wide-association studies for behavioral traits. Psychol Sci. 2014;25:1975-1986.
18. Yang J, Benyamin B, McEvoy BP, *et al*. Common SNPs explain a large proportion of the heritability for human height. Nat Genet. 2010;42:565-569.
19. Visscher PM, Hemani G, Vinkhuyzen AAE, *et al*. Statistical power to detect genetic (co)variance of complex traits using SNP data in unrelated Samples. PLoS Genet. 2014;10:e1004269.
20. St Pourcain B, Skuse DH, Mandy WP, *et al*. Variability in the common genetic architecture of social-communication spectrum phenotypes during childhood and adolescence. Mol Autism. 2014;5:18.
21. Yang L, Neale BM, Liu L, *et al*. Polygenic transmission and complex neuro developmental network for attention deficit hyperactivity disorder: genome-wide association study of both common and rare variants. Am J Med Genet B Neuropsychiatr Genet. 2013;162:419-430.
22. Lee SH, Ripke S, *et al*; Cross-Disorder Group of the Psychiatric Genomics Consortium. Genetic relationship between five psychiatric disorders estimated from genome-wide SNPs. Nature Genet. 2013;45:984-994.
23. Davis LK, Yu D, Keenan CL, *et al*. Partitioning the heritability of Tourette syndrome and obsessive compulsive disorder reveals differences in genetic architecture. PLoS Genet. 2013;9:e1003864.
24. Trzaskowski M, Eley TC, Davis OS, *et al*. First genome-wide association study on anxiety-related behaviours in childhood. PLoS One. 2013;8:e58676.
25. Trzaskowski M, Dale PS, Plomin R. No genetic influence for childhood behavior problems from DNA analysis. J Am Acad Child Adolesc Psychiatry. 2013;52:1048-1056.
26. Conners CK, Sitarenios G, Parker JDA, Epstein JN. The revised Conners' Parent Rating Scale (CPRS-R): factor structure, reliability, and criterion validity. J Abnorm Child Psychol. 1998;26:257-268.
27. Ivanova MY, Achenbach TM, Rescorla LA, *et al*. Preschool psychopathology reported by parents in 23 societies: testing the seven-syndrome model of the Child Behavior Checklist for Ages 1.5-5. J Am Acad Child Psychiatry. 2010;49:1215-1224.
28. Achenbach TM, Rescorla LA. Manual for the ASEBA Preschool Forms and Profiles. Burlington, VT: University of Vermont, Research Center for Children, Youth, and Families; 2000.
29. Achenbach TM, Rescorla LA. Manual for the ASEBA School-Age Forms and Profiles. Burlington, VT: University of Vermont, Research Center for Children, Youth, and Families; 2001.
30. Muratori F, Narzisi A, Tancredi R, *et al*. The CBCL 1.5-5 and the identification of preschoolers with autism in Italy. Epidemiol Psychiatr Sci. 2011;20:329-338.
31. Molins NC, Clopton JR. Teachers' reports of the problem behavior of children in their classrooms. Psychol Rep. 2002;90:157-164.
32. Bartels M, Boomsma DI, Hudziak JJ, van Beijsterveldt TCEM, van den Oord EJCG. Twins and the study of rater (dis)agreement (vol. 12, pg 451, 2007). Psychol Methods. 2008;13:170.
33. Burt SA, McGue M, Krueger RF, Iacono WG. Sources of covariation among the child-externalizing disorders: informant effects and the shared environment. Psychol Med. 2005;35:1133-1144.
34. Tellegen P, Laros JA. Fair assessment of children from cultural minorities: a description of the SON-R non-verbal intelligence tests. Research Institute for Child Psychology and Pathopsychology, and Education Committee, Slovak Commission for UNESCO. Paper presented at the UNESCO seminar: Quality Education for Children from Socially Disadvantaged Settings. Bratislava, Slovakia, June 22-25, 2005.
35. Rietveld M, Dolan C, Van Baal G, Boomsma D. A twin study of differentiation of cognitive abilities in childhood. Behav Genet. 2003;33:367-381.
36. Moore C, O'Keefe SL, Lawhon D, Tellegen P. Concurrent validity of the Snijders-Oomen Nonverbal Intelligence Test 2 1/2-7–Revised with the Wechsler Preschool and Primary Scale of Intelligence–Revised. Psychol Rep. 1998;82:619-625.
37. Bleichrodt N, Drenth PJD, Zaal JN, Resing WCM. Revisie Amsterdamse Kinder Intelligentie Test, RAKIT. Lisse: Swets and Zeitlinger; 1984.
38. Liu EY, Li MY, Wang W, Li Y. MaCH-Admix: genotype imputation for admixed populations. Genet Epidemiol. 2013;37:25-37.
39. Francioli LC, Menelaou A, Pulit SL, *et al*. Whole-genome sequence variation, population structure and demographic history of the Dutch population. Nature Genet. 2014;46:818-825.
40. Yang J, Lee SH, Goddard ME, Visscher PM. GCTA: a tool for genome-wide complex trait analysis. Am J Hum Genet. 2011;88:76-82.
41. van Beijsterveldt CE, Groen-Blokhuis M, Hottenga JJ, *et al*. The Young Netherlands Twin Register (YNTR): longitudinal twin and family studies in over 70,000 children. Twin Res Hum Genet. 2013;16:252-267.
42. Hampel FR, Ronchetti EM, Rousseeuw PJ, Stahel WA. Robust Statistics: the Approach Based on Influence Functions. New York: Wiley; 1986.
43. Lehmann EL. Nonparametrics: Statistical Methods Based on Ranks. New York: Springer; 1975.
44. IBM SPSS Statistics for Windows, Version 21.0 [computer program]. Armonk, NY: IBM Corporation; 2012.
45. St Pourcain B, Whitehouse AJ, Ang WQ, *et al*. Common variation contributes to the genetic architecture of social communication traits. Mol Autism. 2013;4:34.
46. Guey LT, Kravic J, Melander O, *et al*. Power in the phenotypic extremes: a simulation study of power in discovery and replication of rare variants. Genet Epidemiol. 2011;35:236-246.
47. Trzaskowski M, Yang J, Visscher PM, Plomin R. DNA evidence for strong genetic stability and increasing heritability of intelligence from age 7 to 12. Mol Psychiatry. 2014;19:380-384.
48. Haworth CMA, Wright MJ, Luciano M, *et al*. The heritability of general cognitive ability increases linearly from childhood to young adulthood. Mol Psychiatr. 2010;15:1112-1120.
49. Franic S, Dolan CV, van Beijsterveldt CEM, Pol HEH, Bartels M, Boomsma DI. Genetic and environmental stability of intelligence in childhood and adolescence. Twin Res Hum Genet. 2014;17:151-163.
50. Kan KJ, van Beijsterveldt CEM, Bartels M, Boomsma DI. Assessing genetic influences on behavior: informant and context dependency as illustrated by the analysis of attention problems. Behav Genet. 2014;44:326-336.
51. Visscher PM, Goddard ME. A general unified framework to assess the sampling variance of heritability estimates using pedigree or marker-based relationships. Genetics. 2015;199:223-232.
52. Wachs TD. Relation of maternal personality to perceptions of environmental chaos in the home. J Environ Psychol. 2013;34:1-9.
53. Tandon M, Cardeli E, Luby J. Internalizing disorders in early childhood: a review of depressive and anxiety disorders. Child Adolesc Psychiatr Clin N Am. 2009;18:593-610.
54. Hankin BL, Abramson LY, Moffitt TE, Silva PA, McGee R, Angell KE. Development of depression from preadolescence to young adulthood: emerging gender differences in a 10-year longitudinal study. J Abnorm Psychol. 1998;107:128-140.
55. Cicchetti D, Toth SL. The development of depression in children and adolescents. Am Psychol. 1998;53:221-241.
56. Sullivan PF, Daly MJ, Ripke S, *et al*. A mega-analysis of genome-wide association studies for major depressive disorder. Mol Psychiatry. 2013;18:497-511.
57. Asendorpf JB, van de Schoot R, Denissen JJA, Hutteman R. Reducing bias due to systematic attrition in longitudinal studies: the benefits of multiple imputation. Int J Behav Dev. 2014;38:453-460.
58. Tambs K, Ronning T, Prescott CA, *et al*. The Norwegian Institute of Public Health Twin Study of Mental Health: examining recruitment and attrition bias. Twin Res Hum Genet. 2009;12:158-168.
59. Eaton WW, Anthony JC, Tepper S, Dryman A. Psychopathology and attrition in the Epidemiologic Catchment-Area surveys. Am J Epidemiol. 1992;135:1051-1059.

60. Gustavson K, von Soest T, Karevold E, Røysamb E. Attrition and generalizability in longitudinal studies: findings from a 15-year population-based study and a Monte Carlo simulation study. BMC Public Health. 2012;12:918.
61. Wolke D, Waylen A, Samara M, *et al*. Selective drop-out in longitudinal studies and non-biased prediction of behaviour disorders. Br J Psychiatry. 2009;195:249-256.
62. Zhu ZH, Bakshi A, Vinkhuyzen AAE, *et al*. Dominance genetic variation contributes little to the missing heritability for human complex traits. Am J Hum Genet. 2015;96:377-385.
63. Caspi A, Houts RM, Belsky DW, *et al*. The p factor: one general psychopathology factor in the structure of psychiatric disorders? Clin Psychol Sci. 2014;2:119-137.