No Genetic Influence for Childhood Behavior Problems From DNA Analysis

Maciej Trzaskowski, Ph.D., Philip S. Dale, Ph.D., Robert Plomin, Ph.D.

Objective: Twin studies of behavior problems in childhood point to substantial genetic influence. It is now possible to estimate genetic influence using DNA alone in samples of unrelated individuals, not relying on family-based designs such as twins. A linear mixed model, which incorporates DNA microarray data, has confirmed twin results by showing substantial genetic influence for diverse traits in adults. Here we present direct comparisons between twin and DNA heritability estimates for childhood behavior problems as rated by parents, teachers, and children themselves. Method: Behavior problem data from 2,500 UK-representative 12-year-old twin pairs were used in twin analyses; DNA analyses were based on 1 member of the twin pair with genotype data for 1.7 million DNA markers. Diverse behavior problems were assessed, including autistic, depressive, and hyperactive symptoms. Genetic influence from DNA was estimated using genome-wide complex trait analysis (GCTA), and the twin estimates of heritability were based on standard twin model fitting. Results: Behavior problems in childhood—whether rated by parents, teachers, or children themselves—show no significant genetic influence using GCTA, even though twin study estimates of heritability are substantial in the same sample, and even though both GCTA and twin study estimates of genetic influence are substantial for cognitive and anthropometric traits. Conclusions: We suggest that this new type of “missing heritability,” that is, the gap between GCTA and twin study estimates for behavior problems in childhood, is due to nonadditive genetic influence, which will make it more difficult to identify genes responsible for heritability. J. Am. Acad. Child Adolesc. Psychiatry, 2013;52(10):1048–1056. Key Words: behavior problems, cognitive abilities, genome-wide complex trait analysis (GCTA), heritability, twin study

Behavior problems in childhood, such as anxiety, depression, autistic symptoms, hyperactivity, and conduct problems, are common, with a cumulative incidence during childhood of 12% for 1 or more disorders. They are not always transient problems that disappear as children develop: one-half of all lifetime cases of diagnosed psychopathology begin in childhood.2 Furthermore, their heritability is surprisingly high: for example, twin studies using parent ratings typically report heritabilities in the range of 40% for anxiety and depression to 60% for autistic symptoms and hyperactivity.3 Consequently, childhood behavior problems have become the target of genome-wide association (GWA) studies that attempt to identify the genes responsible for their heritability. As in other life sciences,4 these GWA expeditions have come up largely empty-handed.5 This ‘missing heritability’ is the key puzzle in DNA research on complex traits and common disorders—only a small portion of genes responsible for their heritability has been identified.5,6 Although attention has focused on the difficulties in identifying the many genes of small effect responsible for heritability, the other side of the missing heritability gap is the heritability estimate itself, which has relied on family-based studies of twins and adoptees.3 It is now possible to test the validity of these heritability estimates using DNA from unrelated individuals, a method called genome-wide complex trait analysis (GCTA).7,8 GCTA research has shown that the common DNA variants (single-nucleotide polymorphisms [SNPs]) genotyped on DNA arrays used in GWA studies yield...
substantial estimates of genetic influence for height and weight, psychiatric and medical disorders, personality, and cognitive traits. However, these GCTA estimates of genetic influence cannot completely close the heritability gap, in part because GCTA is limited to additive effects of causal variants tagged by the common SNPs on current DNA arrays used in GWA research.

For the first time, we report GCTA estimates for childhood behavior problems as rated by parents, teachers, and the children themselves. To gauge the true breadth of the missing heritability gap, we compared GCTA estimates to twin heritability estimates for the same measures in the same sample. We also compared these results for behavior problems to results for height and weight, and cognitive traits in the same sample.

METHOD
Sample and Genotyping
The sample was drawn from the Twins Early Development Study (TEDS), which is a multivariate longitudinal study that recruited more than 11,000 twin pairs born in England and Wales in 1994, 1995, and 1996. TEDS has been shown to be representative of the UK population. The project received approval from the Institute of Psychiatry ethics committee (05/Q0706/228), and parental consent was obtained before data collection.

The present analyses were limited to children for whom DNA, genome-wide genotyping, and behavior problem and cognitive data were available. Moreover, the twin analyses were based only on twins included in the GCTA analyses, to provide a more precise comparison between GCTA and twin study results.

DNA was available for 3,747 children 11 and 12 years of age (mean age, 11.5 years) whose first language was English and who had no major medical or psychiatric problems. From that sample, DNA samples of 3,665 individuals (only 1 member of a twin pair) were successfully hybridized to Affymetrix GeneChip 6.0 SNP genotyping arrays using standard experimental protocols as part of the WTCCC2 project. In addition to nearly 700,000 genotyped SNPs, more than 1 million other SNPs were imputed from HapMap 2, HapMap 3, and WTCCC controls, using IMPUTE v.2 software. A total of 3,152 DNA samples (from 1,446 males and 1,706 females) survived quality control criteria for ancestry, heterozygosity, relatedness, and hybridization intensity outliers. To control for ancestral stratification, we performed principal component analyses on a subset of 100,000 quality-controlled SNPs after removing SNPs in linkage disequilibrium ($r^2 > 0.2$). Using the Tracy-Widom test, we identified 8 axes with $p < 0.05$, which were used as covariates in GCTA analyses.

Of these 3,152 children, the present analyses were limited to those for whom behavior problem and cognitive data were available. Twin zygosity was diagnosed on the basis of physical similarity, and questionable cases were verified with analysis of DNA markers. As expected, approximately equal numbers of monozygotic (MZ), dizygotic (DZ) same-sex, and DZ opposite-sex twins were included; DZ same-sex and opposite-sex pairs were combined to increase power and because previous twin analyses of these data show no evidence of qualitative or quantitative sex differences in sex-limitation models. For the measures of behavior problems, the numbers of individuals for GCTA analyses range from 2,687 to 2,698 for self-report, 2,687 to 2,700 for parent ratings, and 2,034 to 2,139 for teacher ratings. The numbers of pairs of twins range from 2,668 to 2,683 for self-report, 2,680 to 2,695 for parent ratings, and 1,783 to 1,925 for teacher ratings. The sample sizes for the GCTA results shown are 2,325 for “g” and language, 2,238 for “g”, and mathematics, 2,250 for “g” and reading, and 2,296 for height and weight.

Measures

All of the measures have been reported in previous TEDS publications, which can be consulted for greater detail.

Behavior Problems. The behavior problem measures described in this section have been widely used in the literature. As is the case in the literature, these measures are modestly correlated, 0.33 on average for the scores described below. Given the modest correlation among the measures and the focus of the present paper on comparisons between GCTA and twin estimates for diverse behavior problems, we present results separately for the behavior problem scales rather than conducting multivariate analyses.

Conners (Attention-Deficit/Hyperactivity [ADHD]). ADHD symptoms were assessed via parent-rated questionnaire, which was the DSM-IV–based ADHD scale from the Conners’ Parent Rating Scale–Revised (CPRS-R). The questionnaire consisted of 2 scales: inattentiveness and hyperactivity-impulsivity.

Antisocial Process Screening Device (APS; Psychopathic Symptoms). Antisocial behavior was assessed using parent and teacher ratings on the Antisocial Process Screening Device. The questionnaire included 3 scales: callous-unemotional, impulsivity, and narcissism.

Childhood Asperger Syndrome Test (CAST; Autistic-Like Symptoms). The Childhood Asperger Syndrome Test questionnaire was rated by parents and teachers and includes 3 scales: communication, nonsocial, and social, from which a composite was also formed.

Moods and Feelings Questionnaire (MFQ; Depressive Symptoms). The Moods and Feelings Questionnaire was rated by the children and their parents.
Strengths and Difficulties Questionnaire (SDQ: Behavior Problems). The Strengths and Difficulties Questionnaire was assessed by the children and their parents. The questionnaire includes 4 behavior problem scales (anxiety, conduct, hyperactivity, peer problems) from which a composite was created. The SDQ also includes a positive prosocial scale that was not included in these analyses of behavior problems.

Cognitive Tests. Cognitive data were collected online via the Internet using adaptive branching, which enabled measurement of the full range of ability using a relatively small number of items.

Reading. Four measures of reading were used. Two measures assessed reading comprehension: the reading comprehension subtest of the Peabody Individual Achievement Test (PIAT) and the Global Online Assessment for Learning (GOAL) Formative Assessment in Literacy for Key Stage 3. Reading fluency was assessed by an adaptation of the Woodcock-Johnson III Reading Fluency Test (WJ-RF) and by the Test of Word Reading Efficiency (TOWRE), which was administered by telephone.

Mathematics. Assessment of mathematics targeted 3 components of mathematics: understanding numbers, non-numerical processes, and computation and knowledge. The items for these 3 scales were based on the National Foundation of Educational Research 5-14 Mathematics Series.

Language. Three components of language were assessed: syntax, semantics, and pragmatics. Syntax was measured using the Listening Grammar subtest of the Test of Adolescent and Adult Language. Semantics was assessed using Level 2 of the Figurative Language subtest of the Test of Language Competence. Pragmatics was assessed using Level 2 of the Making Inferences subtest of the Test of Language Competence.

Verbal, Nonverbal, and General Cognitive Abilities. The verbal tests were Wechsler Intelligence Scale for Children as a Processing Instrument (WISC-III-PI) Multiple Choice Information (General Knowledge) and Vocabulary Multiple Choice subtest. The 2 nonverbal reasoning tests were Wechsler Intelligence Scale for Children–III (WISC-III-UK) Picture Completion and Raven’s Standard and Advanced Progressive Matrices. General cognitive ability was indexed as a composite of the 4 verbal and nonverbal tests.

Height and Weight. Height and weight were assessed at age 12 years via self-report.

Composite Measures. To create composite scores for the measures, standardized residuals were derived for each scale regressed on sex and age. Outliers above or below 3 SD from the mean were excluded and the scale was quantile normalized.24,25 The composites were created as unit-weighted means requiring complete data for more than half of the measure’s scales (i.e., 3 of 4 or 2 of 3 scales). All procedures were executed using R (www.r-project.org26).

Statistical Analyses

Genome-Wide Complex Trait Analysis (GCTA). GCTA software was used to conduct these analyses.8 Of note, although GCTA is a software package, for simplicity we will refer to the full process of estimating genetic influence from SNP data simply as GCTA. Before the variance of a trait can be decomposed, the first step is to calculate pairwise genomic similarity between all pairs of individuals in the sample using all genetic markers genotyped or imputed from the SNP array. Because GCTA is designed to estimate genetic variance due to close linkage disequilibrium between unknown causal variants and genotyped SNPs from a sample of unrelated individuals in the population, any close genetic relatedness is eliminated; for this reason, any individual whose genetic similarity is equal to or greater than a third or fourth cousin is removed (estimate of pairwise relatedness > 0.025).

Conceptually, when performing GCTA analysis, we compare a matrix of pairwise genomic similarity to a matrix of pairwise phenotypic similarity using a random-effects mixed linear model.7 In univariate analysis, the variance of a trait can be partitioned using residual maximum likelihood into genetic and residual components. Therefore the residual component includes any source of variance that is not an additive effect of common SNPs, including nonadditive genetic effects, rare variants, environment, gene–environment interaction, and error. Detailed description of this method can be found in Yang et al.7,8 The 8 principal components described earlier were used as covariates. As mentioned in the previous section, all phenotypes were age and sex regressed before analysis.

Twin Analysis. In contrast to GCTA, the twin method models variance/covariance for pairs of related individuals: monozygotic (MZ) twins who are genetically 100% identical, and dizygotic (DZ) twins who share on average 50% of their segregating alleles. Differences in within-pair correlations for MZ and DZ twins are then used to partition the variance into genetic and environmental effects. Variance is attributed to additive genetic influence to the extent that MZ correlations are higher than those for DZ twins. The twin method, unlike GCTA, partitions environmental influence into 2 components: shared or common (C) environmental influences, which is residual MZ twin resemblance not explained by genetics, and nonshared or unique environmental influences (E), which is the extent to which MZ twins differ and includes error of measurement. Detailed description of this A (additive genetics) C (common environment) and E (unique environment; ACE) model and the discussion of related issues can be found elsewhere.3

The twin data were modeled using Cholesky decomposition of the variance within structural equation modeling software OpenMx.27 Standard univariate model-fitting procedures were followed, as described in previous TEDS publications18; standard
errors were derived from 95% CI. In most cases, the influence of shared environment was not significant; however, the full ACE model was used for the comparison with the GCTA results. Although C was not significantly different from zero, dropping it from the model could have inflated A, which would have confounded the comparison between twin and GCTA results. It should be noted that GCTA does not discriminate C and E; both C and E are included in the GCTA estimate of nongenetic residuals. In other words, the “A” and “E” of GCTA and the “A” and “E” of twin ACE model fitting are not the same. For GCTA, “A” denotes additive effects of DNA variants tagged by the common SNPs on our DNA array, and “E” includes all residual variance. In contrast, in twin analysis, “A” includes additive genetic effects of any DNA sequence differences, not just common SNPs; variance not explained by A is partitioned into C and E.

RESULTS

Figure 1 compares GCTA and twin study estimates for the anchor variables of height and weight, as well as for cognitive traits. These results are based on the same individuals and twin pairs used in the present analyses of behavior problems, although the results are highly similar to those previously published for the entire TEDS sample.16 As expected from the literature, the twin study heritability estimates for height and weight are about 80% and the estimates for the cognitive traits are about 50% (~40%–60%). The GCTA estimates are about 40% for height and weight and about 25% (~20%–30%) for the cognitive traits. All of the GCTA estimates are statistically significant, as indicated by the standard errors. These significant and substantial GCTA estimates have 2 important implications. First, they validate the twin method. Second, they imply that sufficiently large GWA studies using current DNA arrays limited to additive effects of common SNPs should be able to account for about 50% of the heritability for height, weight, and cognitive traits. The finding that GCTA estimates are only one-half of the twin heritability estimates is similar to previous reports for these variables and could be due to several factors that either result in GCTA underestimates of twin heritability, such as nonadditive gene–gene interactions, gene–environment interactions, and rare alleles, or to factors that lead to inflation of heritability estimates in twin studies.5

Figure 2 compares GCTA and twin study estimates of heritability for composite measures of behavioral problems for self-report, parent ratings, and teacher ratings. Results for the scales that comprise these composites as well as the full variance decomposition are included in Table S1, available online. Twin heritability estimates are similar to those reported in the literature, that is, about 40% heritability for self-report and about 60% heritability for parent and teacher ratings. In contrast, GCTA estimates are nonsignificant and mostly zero for self-report and parent measures of behavior problems. For teacher ratings, a hint of genetic influence emerged, although these
GCTA estimates of about 10% are not nearly statistically significant, as indicated by the standard errors. The standard errors are larger for GCTA estimates than for twin estimates because GCTA is based on slight (<2.5%) overall pair-by-pair differences in genetic similarity across the 1.7 million SNPs genotyped from the DNA array, whereas the twin estimate is based on the comparison of 100% genetic similarity for MZ twins and 50% similarity for DZ twins for additive genetic effects. However, if the GCTA estimates for behavior problems were one-half of the twin estimates of heritability, as in the case of height and weight and cognitive traits (Figure 1), the GCTA analysis would have adequate power to detect them, as indicated by the standard errors.

**DISCUSSION**

Why do GCTA estimates show no significant genetic influence for diverse childhood behavior problems as rated by parents, teachers, or children themselves, even though twin study estimates of heritability are significant and substantial in the same sample using the same measures, and even though GCTA estimates for cognitive traits are significant and substantial? One broad category of explanations involves mechanisms by which GCTA underestimates twin heritability, more so for behavior problems than for cognitive traits. As mentioned earlier, GCTA underestimates twin heritability because it captures only additive genetic effects tagged by the common SNPs used on GWA arrays. Gene–gene
interactions, gene–environment interactions, and rare alleles will widen the gap between GCTA and twin estimates of heritability. However, it is not clear why this gap would be greater for behavior problems than for cognitive traits.

Nonetheless, we can test 1 of these hypotheses in our study, namely, that nonadditive genetic variance is greater for behavior problems than for cognitive traits. Because MZ twins are identical genetically, they are identical even for interactions among many genes (epistasis), whereas such epistatic effects, on average, scarcely contribute to similarity for DZ twins. Additive genetic effects contribute to MZ twins being twice as similar as DZ twins, whereas the hallmark of nonadditive genetic variance is that MZ twins are more than twice as similar as DZ twins. However, in our study there is no evidence for nonadditive genetic influence either for behavior problems or for cognitive traits: In all cases, MZ correlations are no more than twice as similar as DZ twins. For example, for total behavior problems, MZ and DZ correlations are 0.56 and 0.32, respectively, for self-reports in adolescence; 0.80 and 0.52 for parent ratings; and 0.61 and 0.31 for teacher ratings. Nonetheless, as discussed later, it is possible that nonadditive genetic effects for behavior problems are masked by other factors.

Another, nonmutually exclusive, explanation of why GCTA underestimates twin heritability for behavior problems more than for cognitive abilities is that additive genetic variance might be greater for cognitive traits than for behavior problems. Not previously considered in this context is assortative mating, which increases additive genetic variance for all loci associated with the trait for which spouse correlate. Assortative mating, indexed by the correlation between spouses, is often greater for cognitive traits than for behavior problems: spouse correlations are about 0.40 for cognitive ability. In contrast, assortative mating was reported to be 0.00 for autistic symptoms and 0.02 for hyperactivity in 1 study, and 0.38 and 0.11, respectively, in other studies. Because assortative mating is trait specific and increases additive genetic variance cumulatively over generations, the greater assortative mating for cognitive traits than for behavior problems would increase additive genetic variance for cognitive traits but not for behavior problems. Because GCTA detects only additive genetic variance, assortative mating could thus account for the higher GCTA estimates for cognitive traits than for behavior problems. An obstacle to this hypothesis is that more additive genetic variance for cognitive traits than for behavior problems should lead to higher heritabilities for cognitive traits; however, the results in Figures 1 and 2 show that this is not the case, although this effect could also be masked by countervailing effects, as discussed later.

A second broad category of explanations, and again not mutually exclusive, is that twin studies overestimate heritability for behavior problems more than for cognitive traits. One reason to take this seriously is that twin studies yield higher estimates of heritability than do adoption studies for personality traits, which are related to behavior problems in that personality includes traits such as emotionality, impulsivity, and activity level. Moreover, the first report of GCTA estimates for personality supported the adoption results with estimates of about 10%, the lowest reported GCTA estimates for any domain of behavior before the present study of behavior problems. In contrast, heritability estimates are similar for twin and adoption studies of cognitive traits. It has been suggested that the twin/adoption heritability difference for personality is due to greater nonadditive genetic variance for personality than for cognitive traits, because estimates of heritability from twin studies include nonadditive as well as additive genetic variance (broad heritability), whereas adoption studies that involve first-degree relatives are largely limited to additive genetic variance (narrow heritability). If nonadditive genetic variance is the solution to the twin/adoption heritability difference for personality, it would imply that designs that do not involve MZ twins underestimate heritability caused by nonadditive genetic effects. In other words, twin studies do not overestimate heritability as compared to adoption designs; instead, adoption designs involving first-degree relatives estimate narrow heritability whereas twin studies estimate broad heritability. Although this means that nonadditive genetic variance again emerges as a good candidate for explaining the GCTA/twin heritability gap, this explanation goes against the pattern of MZ and DZ twin correlations for behavior problems in childhood in the present study, which shows no evidence for nonadditive genetic variance, as indicated earlier.

Another methodological possible explanation of the low GCTA heritability estimates could be the skewed distributions, which are often found for measures of behavior problems. As shown in Figure S1, available online, some of the
nontransformed distributions are skewed, but the transformed distributions, which were used in our analyses, are normal. To check on the possibility that the transformation could affect GCTA and twin estimates to different extents, we compared results for both methods using transformed and nontransformed scales and found little difference (see footnote in Figure S1, available online).

Overall, greater nonadditive genetic influence for behavior problems than for cognitive traits emerges as the leading candidate to explain the greater GCTA/twin heritability gap for behavior problems. The only problem with this explanation is that our twin results for behavior problems in children do not indicate nonadditive genetic effects, even though other twin studies of behavior problems in children and twin studies of adult personality point to some nonadditive genetic effects. In the absence of a more parsimonious explanation, we suggest that nonadditive genetic effects for behavior problems in childhood are masked by a general inflation of twin similarity for both MZ and DZ twins. One possibility is that this general twin inflation could be due to experiences that are shared by members of both MZ and DZ twin pairs. However, this possibility seems less implausible, because such shared experiences would seem likely to affect cognitive traits at least as much as behavior problems. A general inflation of twin similarity due to rating bias is another, more promising, possibility: a major difference between behavior problems and cognitive traits is that behavior problems are rated on questionnaires whereas cognitive traits are measured by tests, and ratings are inherently more prone to bias than tests. Such a general inflation of MZ and DZ twin correlations would mask nonadditive genetic variance because it would reduce the difference between MZ and DZ twin correlations. For example, let us suppose that the “true” MZ and DZ twin correlations were 0.3 and 0.1, respectively, suggesting some nonadditive genetic variance. Inflating both twin correlations by 0.1 would result in MZ and DZ correlations of 0.4 and 0.2, respectively, suggesting only additive genetic influence.

Figure 3 illustrates our hypothesis in the context of missing heritability. It introduces a second type of missing heritability. The familiar missing heritability is the extent to which the cumulative effect of all SNPs identified in GWA studies fall short of accounting for twin study heritability estimates. This could be called “missing GWA heritability” to distinguish it from “missing GCTA heritability,” which is the extent to which GCTA estimates fall short of accounting for twin study heritability estimates, and which sets the limit for GWA heritability because both GCTA and GWA are limited to detecting additive effects of common SNPs. We propose that twin studies do not overestimate heritability for behavior problems; rather, twin studies accurately detect nonadditive as well as additive genetic variance for behavior problems, but mostly additive genetic variance for cognitive traits. We do not suggest that all genetic variance for behavior problems is nonadditive, only that there is relatively more nonadditive genetic variance for behavior problems (masked by inflation of twin correlations for both MZ and DZ twins in the present study) and relatively more additive genetic variance for cognitive traits. If the heritability of behavior problems is about 50% and if about one half of the heritability is due to nonadditive genetic variance,
whereas most of the genetic variance for cognitive traits is additive, this would explain the present results, as illustrated in Figure 3. If GCTA heritability for behavior problems were about one half of the GCTA heritability for cognitive traits (e.g., 12% vs. 25%, respectively, in Figure 3), the present study would have little power to detect it, as indicated by the standard errors in Figure 2.

In summary, we propose that a combination of 3 factors is responsible for the greater GCTA/twin heritability gap for behavior problems as compared to cognitive traits. First, nonadditive genetic variance is greater for behavior problems than for cognitive traits, reducing GCTA estimates for behavior problems relative to cognitive traits. Second, greater assortative mating for cognitive traits as compared to behavior problems produces more additive genetic variance for cognitive traits, which results in greater GCTA estimates and thus lowers the GCTA/twin heritability gap for cognitive traits but not for behavior problems. Third, the reason that our twin data do not indicate nonadditive genetic effects for behavior problems in childhood is that these nonadditive genetic effects are masked by inflated correlations for both MZ and DZ twins for ratings of behavior problems. Although this set of hypotheses is speculative, if true, it would suggest that GWA studies will find it more difficult to identify SNPs associated with behavior problems than with cognitive traits. It also suggests that nonadditive genetic variance might contribute importantly to the genetic architecture of behavior problems in childhood, and perhaps in adult personality and psychopathology as well. &

REFERENCES

1. Merikangas KR, He J-P, Burde D, Fisher PW, Boudon K, Koretz DS. Prevalence and treatment of mental disorders among US children in the 2001-2004 NIMHES. Pediatrics. 2010;125:79-81.

2. Kessler RC, Berglund P, Demler O, Jin R, Merikangas KR, Walters EE. Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication. Arch Gen Psychiatry. 2005;62:593-602.

3. Plomin R, DeFries JC, Knopik VS, Neiderhiser JM. Behavioral Genetics. 6th ed. New York: Worth Publishers; 2012.

4. Visscher PM, Brown MA, McCarthy MI, Yang J. Five years of GWAS discovery. Am J Hum Genet. 2012;90:7-24.

5. Plomin R. Child development and molecular genetics: 14 years later. Child Dev. 2012;83:1467-8624.

6. Maher B. Personal genomes: the case of the missing heritability. Nature. 2008;456:18-21.

7. Yang J, Benyamin B, McEvoy BP, Gordon S, Henders AK, Nyholt DR, Madden PA, Heath AC, Martin NG, Montgomery GW, Goddard ME, Visscher PM. Common SNPs explain a large proportion of the heritability for human height. Nat Genet. 2010;42:565-569.

8. Yang JA, Lee SH, Goddard ME, Visscher PM. GCTA: a tool for genome-wide complex trait analysis. Am J Hum Genet. 2011;88:76-82.

9. Yang J, Manolio TA, Pasquale LR, et al. Genome partitioning of genetic variation for complex traits using common SNPs. Nat Genet. 2011;43:519-525.

10. Lee SH, Vray NR, Goddard ME, Visscher PM. Estimating missing heritability for disease from genome-wide association studies. Am J Hum Genet. 2011;88:294-305.

11. Lee SH, DeCandia TR, Ripke S, Yang J. Schizophrenia Psychiatric Genome-Wide Association Study Consortium (PGC-SCZ) International Schizophrenia Consortium (ISC) Molecular Genetics of Schizophrenia Collaboration (MGS) Sullivan PF, Goddard ME, Keller MC, Visscher PM, Wray NR. Estimating the proportion of variation in susceptibility to schizophrenia captured by common SNPs. Nat Genet. 2012;44:247-250.

12. Lubke GH, Hottenga JJ, Walters R, et al. Estimating the genetic variance of major depressive disorder due to all single nucleotide polymorphisms. Biol Psychiatry. 2012;72:707-709.

13. Vinkhuizen AA, Pedersen NL, Yang J, et al. Common SNPs explain some of the variation in the personality dimensions of neuroticism and extraversion. Transl Psychiatry. 2012;2:e102.

14. Davies G, Tenea A, Payton A, et al. Genome-wide association studies establish that human intelligence is highly heritable and polygenic. Mol Psychiatry. 2011;16:996-1005.

15. Deary IJ, Yang J, Davies G, et al. Genetic contributions to stability and change in intelligence from childhood to old age. Nature. 2012;482:212-215.

16. Plomin R, Haworth CMA, Mathur BM, Price T, 2 WTCCC, Davis OSP. Common DNA markers can account for more than half of the genetic influence on cognitive abilities. Psychol Sci. 2013;24:562-568.

17. Oliver BR, Plomin R. Twins Early Development Study (TEDS): a multivariate, longitudinal genetic investigation of language, cognition and behaviour problems from childhood through adolescence. Twin Res Hum Genet. 2007;10:96-105.

18. Kovas Y, Haworth CMA, Dale PS, Plomin R. The genetic and environmental origins of learning abilities and disabilities in the early school years. Monogr Soc Res Child Dev. 2007;72:1-144.

19. Howie BN, Donnelly P, Marchini J. A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. PLoS Genet. 2009;5:e1000529.

Accepted July 31, 2013.

Drs. Trzaskowski and Plomin are with the Social, Genetic, and Developmental Psychiatry Centre, Institute of Psychiatry, King’s College London. Dr. Dale is with the University of New Mexico.

The Twins Early Development Study (TEDS) is supported by a program grant to Dr. Plomin from the UK Medical Research Council (G0901245, and previously G0500079), with additional support from the US National Institutes of Health (HD044454, HD059215). Genome-wide genotyping was made possible by a grant from the Wellcome Trust to the Wellcome Trust Case Control Consortium 2 project (085475/B/08/Z, 085475/Z/08/Z). Dr. Plomin is supported by a Medical Research Council Research Professorship award (G19/2) and a European Advanced Investigator award (295326). Mr. Trzaskowski is supported by a Medical Research Council studentship.

Dr. Trzaskowski served as the statistical expert for this research. He had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

The authors thank the reviewers for highlighting several valuable issues.

Disclosure: Drs. Dale, Plomin, and Trzaskowski report no biomedical financial interests or potential conflicts of interest.

Correspondence to Maciej Trzaskowski, Box PO80, SGDP Centre, Institute of Psychiatry, DeCrespigny Park, London SE5 8AF UK; email: maciej.trzaskowski@kcl.ac.uk

08908567/$36.00/©2013 American Academy of Child and Adolescent Psychiatry

http://dx.doi.org/10.1016/j.jaac.2013.07.016
20. Fellay J, Shianna KV, Ge D, et al. A whole-genome association study of major determinants for host control of HIV-1. Science. 2007;317:944-947.
21. Patterson N, Price AL, Reich D. Population structure and eigenanalysis. PLoS Genet. 2006;2:2074-2093.
22. Greven C, Rijsdijk F, Plomin R. A twin study of ADHD symptoms in early adolescence: hyperactivity-impulsivity and inattentiveness show substantial genetic overlap but also genetic specificity. J Abnorm Child Psychol. 2011;39:265-275.
23. Haworth CMA, Davis OSP, Plomin R. Twins Early Development Study (TEDS): a genetically sensitive investigation of cognitive and behavioural development from childhood to young adulthood. Twin Res Hum Genet. 2012;16:117-125.
24. van der Waerden BL. On the sources of my book “Moderne Algebra”. Historia Mathematica. 1975;2:31-40.
25. Lehmann EL. Nonparametrics: Statistical Methods Based on Ranks. San Francisco: Holden-Day; 1975.
26. R Core Development Team. R: A Language and Environment for Statistical Computing (version 2.13.1). 2011. Available at: http://www.R-project.org/. Accessed November 3, 2012.
27. Boker S, Neale M, Maes H, et al. OpenMx: An open source extended structural equation modeling framework. Psychometrika. 2011;76:306-317.
28. Jensen AR. Genetic and behavioural effects of nonrandom mating. In: Osbourne RT, Noble CE, Weyl N, eds. Human Variation: The Biopsychology of Age, Race, and Sex. New York: Academic Press; 1978:51-105.
29. van Steijn DJ, Richards JS, Oerlemans AM, et al. The co-occurrence of autism spectrum disorder and attention-deficit/hyperactivity disorder symptoms in parents of children with ASD or ASD with ADHD. J Child Psychol Psychiatry. 2012;53:954-963.
30. Constantino JN, Todd RD. Intergenerational transmission of subthreshold autistic traits in the general population. Biol Psychiatry. 2005;57:655-660.
31. Boomsma DI, Saviouk V, Hottenga JJ, et al. Genetic epidemiology of attention deficit hyperactivity disorder (ADHD index) in adults. PLoS One. 2010;5:e10621.
32. Plomin R, Corley R, Caspi A, Fulker DW, DeFries JC. Adoption results for self-reported personality: evidence for nonadditive genetic effects? J Pers Soc Psychol. 1998;75:211-218.
FIGURE S1  (A) Histograms of untransformed composite scales, (B) Histograms of composite quantile normalized scales. Note: As seen in Figure S1A, some of the composite scales are skewed, as is typical of behavior problem scales. However, Figure S1B shows that van der Wearden transformation (see Method) normalizes the scales. It is noteworthy that despite the considerable transformation of the distributions, the correlation between the untransformed and transformed scales are high (0.90, 0.96, 0.96, 0.99, 0.92, 0.98, 0.98, 0.97, 0.97, 0.94, 0.95, 0.94, 0.95, 0.95, respectively), indicating that the transformation did not drastically disrupt the rank-order structure of the data. In the Results section of the text, we presented results for the transformed scales; however, as a further check on the effect of non-normality, we also compared genome-wide complex trait analysis (GCTA) and twin point estimates of heritability for the transformed and untransformed scales. We found that the largest GCTA heritability difference was 0.04 and the largest twin heritability difference was also 0.04. ‘lcmfq1’ = child self-rated Moods and Feelings Questionnaire (MFQ); ‘lcsanxt1’ = child self-rated Strengths and Difficulties Questionnaire (SDQ) Anxiety; ‘lcscont1’ = child self-rated SDQ conduct; ‘lcshyp1’ = child self-rated SDQ hyperactivity; ‘lcspert1’ = child self-rated SDQ peer problems; ‘lcsbeht1’ = child self-rated SDQ composite; ‘lpmfq1’ = parent-rated MFQ composite; ‘lpsbeht1’ = parent-rated SDQ composite; ‘ltapsdt1’ = teacher-rated Antisocial Process Screening Device (APSD) composite; ‘ltcastt1’ = teacher-rated Childhood Asperger Syndrome Test (CAST) composite; ‘ltsbeht1’ = teacher-rated SDQ composite; ‘lpconnt1’ = parent-rated Conners attention-deficit/hyperactivity disorder (ADHD) composite.
|          | Twin Heritability       | GCTA Heritability       |
|----------|-------------------------|-------------------------|
|          | A  | SE | C  | SE | E  | SE | n/Pairs | V(G) | SE | V(e) | SE | n   |
| MFQ (depressive symptoms) | 0.38 | 0.06 | 0.09 | 0.05 | 0.53 | 0.02 | 2,683 | 0.00 | 0.12 | 1.00 | 0.12 | 2,698 |
| SDQ Behavior Problems composite | 0.44 | 0.06 | 0.09 | 0.05 | 0.47 | 0.02 | 2,668 | 0.00 | 0.12 | 1.00 | 0.13 | 2,690 |
| Anxiety  | 0.41 | 0.03 | 0.00 | 0.01 | 0.59 | 0.02 | 2,668 | 0.02 | 0.12 | 0.99 | 0.13 | 2,687 |
| Conduct  | 0.37 | 0.06 | 0.06 | 0.04 | 0.57 | 0.02 | 2,670 | 0.00 | 0.12 | 1.00 | 0.12 | 2,690 |
| Hyperactivity | 0.46 | 0.03 | 0.00 | 0.01 | 0.54 | 0.02 | 2,672 | 0.00 | 0.12 | 1.00 | 0.12 | 2,687 |
| Peer problems | 0.40 | 0.05 | 0.02 | 0.03 | 0.57 | 0.02 | 2,674 | 0.00 | 0.12 | 1.00 | 0.12 | 2,692 |

**Note:** Composites are also presented in Figure 2. A = heritability estimate; ADHD = attention-deficit/hyperactivity disorder; APSD = Antisocial Process Screening Device; C = shared environment estimate; CAST = Childhood Asperger Syndrome Test; E = unique environment estimate; MFQ = Moods and Feelings Questionnaire; n = number of unrelated individuals with available data; n/Pairs = number of twin pairs with available data; SDQ = Strengths and Difficulties Questionnaire; SE = standard error; V(G) = variance explained by additive genetic factors.