Different heritabilities but shared etiological influences for parent, teacher and self-ratings of ADHD symptoms: an adolescent twin study

A. Merwood[1,*], **C. U. Greven**[1,2], **T. S. Price**[1,3], **F. Rijsdijk**[1], **J. Kuntsi**[1], **G. McLoughlin**[1], **H. Larsson**[4,5], and **P. J. Asherson**[1]

[1]King's College London, MRC Social, Genetic and Developmental Psychiatry Centre, Institute of Psychiatry, London, UK [2]Radboud University Nijmegen Medical Center, Donders Institute for Brain Cognition and Behavior, The Netherlands [3]University of Pennsylvania, Perelman School of Medicine, Philadelphia, PA, USA [4]Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden [5]Centre of Neurodevelopmental Disorders, Karolinska Institutet, Stockholm, Sweden

Abstract

Background—Parent and teacher ratings of attention deficit hyperactivity disorder (ADHD) symptoms yield high estimates of heritability whereas self-ratings typically yield lower estimates. To understand why, the present study examined the etiological overlap between parent, teacher and self-ratings of ADHD symptoms in a population-based sample of 11–12-year-old twins.

Method—Participants were from the Twins Early Development Study (TEDS). ADHD symptoms were assessed using the Strengths and Difficulties Questionnaire (SDQ) hyperactivity scale completed by parents, teachers and children. Structural equation modeling was used to examine genetic and environmental contributions to phenotypic variance/covariance.

Results—The broad-sense heritability of ADHD symptoms was 82% for parent ratings, 60% for teacher ratings and 48% for self-ratings. *Post-hoc* analyses revealed significantly higher heritability for same-teacher than different-teacher ratings of ADHD (76% *v*. 49%). A common pathway model best explained the relationship between different informant ratings, with common genetic influences accounting for 84% of the covariance between parent, teacher and self-rated ADHD symptoms. The remaining variance was explained by rater-specific genetic and non-shared environmental influences.

Conclusions—Despite different heritabilities, there were shared genetic influences for parent, teacher and self-ratings of ADHD symptoms, indicating that different informants rated some of the same aspects of behavior. The low heritability estimated for self-ratings and different-teacher ratings may reflect increased measurement error when different informants rate each twin from a pair, and/or greater non-shared environmental influences. Future studies into the genetic

*Address for correspondence: Mr A. Merwood, MRC SGDP Centre, PO80, Institute of Psychiatry, De Crespigny Park, Denmark Hill, London SE5 8AF, UK. (andrew.merwood@kcl.ac.uk) .

influences on ADHD should incorporate informant data in addition to self-ratings to capture a pervasive, heritable component of ADHD symptomatology.

Keywords

ADHD; etiology; genetics; rater effects; twin study

Introduction

Attention deficit hyperactivity disorder (ADHD) is characterized by developmentally inappropriate and impairing symptoms that aggregate into two dimensions of inattention and hyperactivity-impulsivity (APA, 1994). ADHD is one of the most common childhood psychiatric diagnoses and frequently persists into adulthood (Faraone *et al.* 2006). Research suggests that ADHD symptoms are distributed continuously throughout the population, with clinical cases derived from extreme symptom scores (Chen *et al.* 2008; Larsson *et al.* 2012*a*).

The methods used to assess ADHD symptoms vary throughout the lifespan. In childhood, symptoms are usually rated by parents and teachers; in late adolescence and adulthood, symptoms are more frequently self-rated (Asherson, 2005). Parent, teacher and self-ratings of ADHD symptoms correlate only moderately, around 0.3 to 0.5 (Achenbach & Rescorla, 2000; Goodman, 2001).

Univariate twin studies suggest that heritability estimates for ADHD symptoms are to some extent informant specific. Parent and teacher ratings of childhood and adolescent ADHD symptom scores typically yield high heritability estimates (70–80%) (Nikolas & Burt, 2010) whereas studies that use self-ratings consistently estimate lower heritability (<50%). This is true of self-ratings obtained in adolescence (Young *et al.* 2000; Martin *et al.* 2002; Ehringer *et al.* 2006) and of retrospective and current self-ratings obtained in adulthood (Schultz *et al.* 2006; Van Den Berg *et al.* 2006; Haberstick *et al.* 2008; Boomsma *et al.* 2010; Larsson *et al.* 2012*b*). Some studies also estimate lower heritability when different teachers rate each twin from a pair, rather than the same teacher rating each twin (Simonoff *et al.* 1998; Saudino *et al.* 2005; Derks *et al.* 2006; Hartman *et al.* 2007).

One explanation for the lower heritability of self-ratings is that they may be less reliable than other informant ratings of ADHD symptoms. Low reliability leads to lower identical (monozygotic, MZ) within-twin correlations, imposing a ceiling on heritability estimates by increasing measurement error (Plomin *et al.* 2008). Indeed, measurement error has been proposed as an explanation for the lower heritability estimated for different-teacher ratings of ADHD (Hartman *et al.* 2007).

Parent ratings of ADHD symptoms often show non-additive in addition to additive genetic influences (Burt, 2009), whereas teacher and self-ratings tend to show only additive genetic influences. This may reflect a rater contrast effect, whereby parents contrast the behavior of their twins and underestimate the similarity of non-identical (dizygotic, DZ) twins (Simonoff *et al.* 1998; Wood *et al.* 2010). In genetic modeling, contrast effects and genetic non-additivity both lead to low within-twin correlations for DZ twins. However, contrast effects can be distinguished from genetic non-additivity by greater variance in the behaviors of DZ than MZ twins (Price *et al.* 2005).

Because of these nuances, an important question is whether different informants rate the same aspects of ADHD-related behaviors. Rater differences can occur because of genuine differences in perspective and/or rater biases (Derks *et al.* 2006) and can be disentangled

through multivariate twin studies that use multiple informant data: unique genetic influences indicate that different informants rate unique but valid aspects of behavior; unique environmental influences may reflect rater-specific bias or measurement error; and overlapping genetic influences and overlapping environmental influences indicate the extent to which different informants rate the same aspects of behavior (Hewitt *et al.* 1992).

Bivariate twin studies have identified both common and unique genetic influences on parent and teacher ratings of ADHD symptoms. This suggests that the same aspects of behaviour are rated by different informants, in addition to unique aspects of behavior (Simonoff *et al.* 1998; Thapar *et al.* 2000; Martin *et al.* 2002; Nadder *et al.* 2002; Derks *et al.* 2006; Hartman *et al.* 2007; McLoughlin *et al.* 2011). However, there are as yet no studies that have investigated the relationship between parent and teacher ratings and self-ratings of ADHD symptoms.

In the current study we used a large population-based sample of 11–12-year-old twins to investigate the genetic and environmental contributions to individual differences in parent, teacher and self-ratings of ADHD symptoms. Multivariate genetic modeling was used to evaluate the extent to which the different informant ratings reflect the same and/or specific views of ADHD. Characterizing the phenotypic and etiological relationships between self and other informant ratings of ADHD is particularly relevant to our understanding of the developmental course of ADHD, not least because self-ratings are increasingly relied upon in the transition from adolescence toadulthood but also because etiological research depends on the quality of ratings.

Method

Sample and procedure

Participants were from the Twins Early Development Study (TEDS), a population-representative sample of all twin birth records in the UK for the years 1994 to 1996 (Oliver & Plomin, 2007). Ethical approval was provided by the Research Ethics Committee of the Institute of Psychiatry, King's College London. Twin zygosity was initially determined by parental report using a questionnaire with 95% accuracy (Price *et al.* 2000) and was subsequently verified using DNA obtained from cheek swabs (Oliver & Plomin, 2007).

Data were collected when twins were aged 11–12 years. Exclusion criteria were severe medical problems at the time of assessment, severe problems at birth or during pregnancy, and unknown or uncertain sex or zygosity. Pairs in which ADHD symptom ratings were unavailable for either twin were also excluded. The final sample thus comprised 6372 twin pairs: parent ratings of ADHD were available for 5590 pairs (including two incomplete pairs); teacher ratings for 5217 pairs (including 1069 incomplete pairs); and self-ratings for 5621 pairs (including 84 incomplete pairs). The number of pairs is presented by sex, zygosity and informant in Table 1. The mean age of participating twins was 11.28 (s.d.=0.70) years.

Measures

The Strengths and Difficulties Questionnaire (SDQ; Goodman, 2001)—The SDQ is a 25-item questionnaire designed to measure common mental health problems during childhood and adolescence. ADHD symptoms were assessed using the SDQ hyperactivity scale, a five-item measure of inattention ('easily distracted, concentration wanders'), hyperactivity ('constantly fidgeting or squirming') and impulsivity ('thinks things out before acting'). Ratings were made using a three-point Likert scale. There are insufficient items to provide a valid separation of the inattentive and hyperactive/impulsive symptoms into

separate subscales and the loading of all five items onto a hyperactivity scale is supported by factor analysis (Goodman, 2001; Van Roy *et al.* 2008). As in previous research using the SDQ hyperactivity scale (Price *et al.* 2005), scores across all five items were averaged to create a total ADHD symptom score. The scale was completed by parents and teachers, and was self-rated by children. Cronbach's was 0.76 for parent ratings, 0.86 for teacher ratings and 0.69 for self-ratings.

Analyses

The twin method (Plomin *et al.* 2008) was used to decompose phenotypic variance/covariance into additive genetic (A), non-additive genetic (D) and non-shared environmental (E) components. Broad-sense heritability estimates were derived from the sum of A+D. Measurement error was accounted for by the component E. Genetic modeling was conducted using the structural equation modeling program Mx (Neale, 1997), which estimated genetic and environmental correlations within and across twin pairs. It is assumed that both additive genetic (r_A) and non-additive genetic (r_D) correlations within MZ twin pairs are 1.00 because 100% of genetic variation is shared. Within DZ pairs, r_A is assumed to be 0.50 and r_D is assumed to be 0.25, reflecting on average 50% additive genetic similarity and 25% non-additive genetic similarity. The non-shared environment (E) is unique to individuals and therefore correlates at zero within MZ and DZ twin pairs.

Prior to genetic modeling, raw data were square-root transformed to correct for non-normal distribution and were regressed to correct for the effects of age and sex, a standard twin modeling procedure (McGue & Bouchard, 1984). All transformed/regressed variables showed approximately normal distribution (skewness and kurtosis within range ±1). Mx used full-information maximum likelihood estimation, in which a likelihood statistic (−2 log likelihood or −2LL) of the data for each observation was calculated. Likelihood-based confidence intervals (CIs) were used to assess the accuracy and significance of parameter estimates (Neale & Miller, 1997).

Univariate genetic modeling—Univariate genetic models decomposed the variance in parent, teacher and self-ratings of ADHD symptoms into the components ADE. Models including contrast effects (*b*) were fit when low DZ within-twin correlations were observed in the presence of greater variances for DZ than MZ twins. ADE and ADE+*b* models were tested separately as this provides the greatest power to detect genetic non-additivity (Rietveld *et al.* 2003).

Full sex-limitation models were used to test whether the genetic and environmental factors influencing males were different to those influencing females (qualitative sex differences), whether the magnitude of factor loadings influencing males and females were different (quantitative sex differences), and whether there were differences in phenotypic variances between males and females (scalar sex differences). The full sex-limitation model (1) contains three nested submodels (2–4) and can be explained as follows (Davis *et al.* 2008):

1. The full sex-limitation model allowed quantitative and qualitative differences in the parameter estimates between males and females, and freely estimated either r_A or r_D.

2. The common effects sex-limitation model allowed quantitative sex differences between males and females but no qualitative differences, fixing r_A to 0.5 and r_D to 0.25 in the DZ opposite-sex group.

3. The scalar effects sex-limitation model allowed variance differences between males and females but no qualitative or quantitative differences, fixing r_A to 0.5 and r_D to 0.25 in the DZ opposite-sex group and constraining the variance components for

males to be a scalar multiple of female variance components. (To test for sex differences in contrast effects for this study, parameter estimates for *b* were equated for males and females as an additional step.)

4. The null model equated all parameter estimates for males and females, testing the hypothesis that there were no sex differences.

The relative fit of nested models was assessed using the likelihood ratio test (LRT), calculated by comparing the difference in –2LL against a x^2 distribution with degrees of freedom equal to the number of parameters eliminated in the reduced model. A significant result indicated deterioration in model fit. Goodness of fit was also determined using the Bayesian information criterion (BIC) statistic, which favors parsimony in large sample sizes (Raftery, 1995). Lower BIC values indicated better model fit; differences of more than 10 identified a strong preference for the model with the lower BIC value.

Multivariate genetic modeling—Multivariate genetic models were used to examine covariance between parent, teacher and self-ratings of ADHD symptoms. These used cross-twin cross-trait (CTCT) correlations to decompose phenotypic covariation into genetic and environmental components. Contrast effects were included where appropriate, based on the univariate genetic results. The fit of different multivariate models was compared using the BIC statistic, with the decision on whether to accept reduced models based on the LRT. Three classes of model were tested:

1. Cholesky decomposition. We interpreted the mathematically equivalent correlated factors solution (Fig. 1*a*) (Loehlin, 1996). It parameterized the extent to which influences underlying one informant rating of ADHD symptoms (e.g. parent) also influenced other ratings (e.g. teacher, self). Each rating of ADHD symptoms was decomposed into its genetic and environmental components (ADE) and the correlation of these components was estimated across informants.

2. Independent pathway model (Fig. 1*b*). This was based on a biometric model, in which common variance components (ADE) loaded onto the ADHD symptom ratings to account for phenotypic covariance. These represented genetic and environmental influences that contributed to all informant ratings. Residual components (ade) accounted for the remaining variance that was specific to each informant.

3. Common pathway model (Fig. 1*c*). This was based on a psychometric model, in which common components (ADE) loaded onto a latent ADHD factor with variance constrained to 1.00. The latent factor accounted for covariance among the different informant ratings, representing a common, pervasive view of ADHD. Residual components (ade) accounted for the remaining variance that was specific to each informant.

Results

Descriptive statistics and correlations are presented in Table 1. Tests of mean differences were performed on the raw data using robust regressions in Stata (StataCorp, 2007) to control for dependence in the observations from twin pairs (Williams, 2000). Mean ADHD symptom scores were significantly higher for males than females based on ratings from parents (t=22.24, p<0.001), teachers (t=25.20, p<0.001) and children (t=17.00, p<0.001).

Twin variances and correlations were estimated using a saturated model fit to the transformed/regressed data. In this model phenotypic correlations were constrained to be equal regardless of sex or zygosity whereas variances, within-twin and CTCT correlations

were estimated separately for the different sex-by-zygosity groups. To test for differences, variances were equated across sex-by-zygosity groups to see whether this resulted in a significant deterioration in model fit based on the LRT. Variances were significantly higher for males than females based on ratings from parents (x^2=28.68, p<0.001) and teachers (x^2=200.54, p<0.001), indicating probable scalar sex differences. For parent ratings variances were also significantly higher for DZ than MZ twins (x^2=18.72, p=0.001), indicating possible contrast effects.

For parent ratings the DZ within-twin correlations were less than half the MZ correlations, further suggesting contrast effects and/or non-additive genetic influences on phenotypic variance. For teacher ratings the DZ correlations were roughly half the MZ correlations, suggesting additive genetic influences. For self-ratings the DZ correlations were less than half the MZ correlations, suggesting some non-additive genetic influences. CTCT correlations for the DZ pairs were less than half of those for the MZ pairs, suggesting non-additive genetic influences on phenotypic covariance. Phenotypic correlations were 0.34 (95% CI 0.32–0.36) for parent with teacher ratings, 0.45 (95% CI 0.45–0.47) for parent with self-ratings and 0.29 (95% CI 0.27–0.31) for teacher with self-ratings.

Univariate genetic results

Full sex-limitation models indicated variance sex differences for all informant ratings of ADHD symptoms. For parent ratings, the most parsimonious model was an AE scalar model, with a contrast effect (*b*) that was equated for males and females (A^2=0.82, 95% CI 0.80–0.83; *b*=–0.04, 95% CI –0.05 to –0.03). The most parsimonious models were an AE scalar model for teacher ratings (A^2=0.60, 95% CI 0.58–0.63) and an ADE scalar model for self-ratings (A^2=0.28, 95% CI 0.15–0.41; D^2=0.20, 95% CI 0.06–0.34). Broad-sense heritability estimates were thus 82%, 60% and 48% respectively. Fit statistics for all univariate models tested are presented in the online supplementary material (Table S1).

Multivariate genetic results

Based on the univariate results the multivariate models included a scalar to account for variance sex differences for all informant ratings of ADHD, in addition to a contrast effect (*b*) for parent ratings only. The BIC statistic indicated a strong preference for the common pathway model, from which a reduced model parameterizing ADE at the common level, ae at the residual level and *b* for parent ratings provided the most parsimonious fit. Parameter estimates for this model are presented in Table 2. Fit statistics for all multivariate models tested are presented in the online supplementary material (Table S2).

A common factor accounted for similarities among the different informant ratings of ADHD symptoms. This factor was highly heritable (A^2+D^2=0.84), with the remaining variance explained by the non-shared environment. When examining loadings of each informant rating onto the latent factor, it was apparent that common genetic influences (common A+D) accounted for 43% of the total variance in parent ratings, 17% in teacher ratings and 32% in self-ratings (see footnote of Table 2 for calculations). These results indicate that parent, teacher and self-ratings assessed some of the same aspects of ADHD-related behavior, and that common genetic influences accounted for most of the similarity between informants.

The remaining variance for each informant rating was accounted for by residual genetic and environmental factors. The presence of residual genetic influences indicate that all informants rated unique but valid aspects of ADHD-related behavior whereas residual non-shared environmental influences indicate that different informant reports were also influenced by the unique environment and/or measurement error.

Post-hoc analyses of same/different-teacher ratings

In genetic modeling the heritability estimated for teacher ratings was lower than expected. Previous research indicates that this can occur when same- and different-teacher ratings of ADHD symptoms are combined (Derks *et al.* 2006). We therefore split the sample based on whether both twins from a pair had the same teacher (*n*=1868 pairs) or different teachers (*n*=3349 pairs) at school and repeated all genetic modeling separately for these groups.

We first conducted univariate modeling. For both groups, the most parsimonious models were AE scalar models. These estimated higher heritability for same-teacher ratings (A^2=0.76, 95% CI 0.73–0.78)] than different-teacher ratings (A^2=0.49, 95% CI 0.44–0.53). Non-overlapping CIs indicated that this was a significant difference. Fit statistics are presented in the online supplementary material (Table S3).

We then refit the common pathway model. For both groups a model that parameterized ADE at the common level and ae at the residual level provided the best fit. The model for the different-teacher group also incorporated a contrast effect (*b*) for parent-rated ADHD symptoms; however, in the same-teacher group the contrast effect for parent ratings was non-significant and could be removed in the interests of model parsimony. Additive genetic influences on the latent factor for the same-teacher group were also non-significant but were retained in the model because it is considered biologically implausible to model genetic non-additivity in the absence of additive genetic effects. Non-significance of these parameter estimates probably reflects the smaller sample size of the same-teacher group.

In both the same-teacher and different-teacher models, a highly heritable latent factor accounted for covariance among parent, teacher and self-ratings of ADHD symptoms (A^2+D^2=0.85 and 0.83 respectively). This is consistent with results reported for the whole sample. Residual genetic influences (a^2) were significantly higher in the same-teacher than the different-teacher models. Parameter estimates are presented in Table 3 and model fit statistics in the online supplementary Table S4.

Discussion

This study investigated the etiological relationship between parent, teacher and self-ratings of ADHD symptoms. There were two main findings. First, heritability estimates were lower for self-ratings (48%) than for parent (82%) or teacher (60%) ratings, even though all ratings were obtained concurrently during early adolescence. Second, multivariate modeling indicated shared and unique etiological influences on different informant ratings, suggesting shared but also rater-specific views of ADHD-related behaviors.

Previous twin studies of self-rated ADHD symptoms have reported univariate heritabilities below 50% in adolescence and adulthood (Young *et al.* 2000; Martin *et al.* 2002; Ehringer *et al.* 2006; Schultz *et al.* 2006; Van Den Berg *et al.* 2006; Haberstick *et al.* 2008; Boomsma *et al.* 2010; Larsson *et al.* 2012*b*). In the current study we extended these findings to a younger age group, finding similar heritability (48%) for self-ratings of ADHD symptoms in 11–12-year-old twins. This focus on early adolescence indicates that the lower heritability associated with self-ratings is not exclusive to late adolescence and adulthood, and challenges the conclusion that ADHD might be a less heritable phenotype in adults (Boomsma *et al.* 2010; Saviouk *et al.* 2011).

As expected from a recent meta-analysis (Nikolas & Burt, 2010), the heritability estimate for parent ratings was high (82%), but was lower than expected for teacher ratings (60%). When we divided our data into two samples, based on whether the behaviors for both twins from a pair were rated by the same or different teachers, we estimated a significantly higher

heritability from same-teacher ratings (76% *v.* 49%). This observation has been reported previously (Simonoff *et al.* 1998; Saudino *et al.* 2005; Derks *et al.* 2006; Hartman *et al.* 2007) and therefore seems to be a robust finding.

It is noteworthy that the heritability estimates derived from same-teacher ratings were similar to parent ratings whereas the estimates from different-teacher ratings were similar to self-ratings. This suggests that having a single informant rate the behaviors of both twins from a pair (either a parent or the same teacher) leads to higher heritability estimates than having ratings by different informants for each twin (either the children themselves or different teachers). There are several possible conclusions.

One conclusion is that the different-informant ratings may be more sensitive to genuine non-shared environmental influences on behavior, such as peer relationships or teacher characteristics. If this is the case, then different-informant ratings may provide more accurate heritability estimates that better account for non-shared environmental effects. Another conclusion is that of gene–environment interaction, which occurs when genetic influences depend on the environment. This was the conclusion of a recent twin study that suggested that exposure to different teachers and the corresponding classroom environments triggered different externalized behaviors in each twin from a pair (Lamb *et al.* 2012). A third conclusion is that different-informant ratings may be associated with increased measurement error, a likely scenario because reliability between ratings will always be lower when two raters rather than just one is involved (unless inter-rater reliability approaches 1). If this is the case then the different-informant ratings may underestimate heritability. Unfortunately, we were unable to distinguish genuine non-shared environmental effects from measurement error in this study, so cannot say which of these explanations may be correct.

An additional explanation that must be considered in relation to the low heritability of self-ratings is that children may be unreliable informants of their own behavior. Previous research has shown that the SDQ hyperactivity scale is less reliable for self-ratings than parent or teacher ratings, based on internal consistency and retest stability in children and adolescents (Goodman, 2001). Moreover, the internal consistency of self-ratings from the SDQ hyperactivity scale is found to increase with age, from 10–13 years (α=0.57) to 13–16 years (α=0.65) and 16–19 years (α=0.66) (Van Roy *et al.* 2008). Children may therefore be less reliable informants than older individuals when rating their own ADHD symptoms. In the present study the internal consistency for self-ratings was acceptable (α=0.69), although not as good as for parent (α=0.76) or teacher (α=0.86) ratings. Nonetheless, this suggests that the children who participated in this study were reasonably reliable when assessing their own ADHD symptomatology.

In the multivariate genetic modeling a highly heritable latent factor accounted for similarity between parent, teacher and self-ratings of ADHD symptoms, indicating that the overlap between different informant ratings was largely due to a common set of genetic effects. *Post-hoc* analyses showed similar results when same-teacher and different-teacher ratings were considered separately. However, the loading of teacher ratings onto the latent factor was always significantly lower than the loadings of parent or self-ratings, indicating that the greatest similarity was between the parents and children. The weaker association of teacher ratings with this pervasive view is in line with previous studies showing distinct and/or shared etiological influences for parent and teacher ratings of ADHD symptoms (Simonoff *et al.* 1998; Thapar *et al.* 2000; Martin *et al.* 2002; Nadder *et al.* 2002; Derks *et al.* 2006; Hartman *et al.* 2007; McLoughlin *et al.* 2011). Because of this, and because of the finding of residual genetic influences on parent, teacher and self-ratings, rater-specific effects are likely

to be valid indicators of different aspects of ADHD-related behaviors, perhaps reflecting differences at home and at school.

Finally, we can comment on the role of contrast effects and genetic non-additivity across different informant ratings of ADHD. Consistent with previous research using the SDQ (Price *et al.* 2001, 2005; Saudino *et al.* 2005), univariate modeling identified significant contrast effects for parent ratings only. Conversely, there were significant non-additive genetic influences on self-ratings, a finding not reported previously. The multivariate model also included non-additive genetic influences on the common factor, indicating that these were important with regard to the overlap between informants.

The results should be interpreted in the context of several limitations. First, we examined ADHD symptoms in a population-based twin sample, meaning that the findings may not generalize to clinical cases of ADHD. Second, we used a short five-item measure of ADHD symptoms (the SDQ hyperactivity scale) rather than an 18-item questionnaire. We took this approach because self-ratings on more comprehensive measures of ADHD symptoms were unavailable; however, the scale has been used to assess ADHD symptoms in previous twin studies in this sample (Price *et al.* 2001, 2005; Saudino *et al.* 2005). Third, because we used the SDQ we were unable to examine the dimensions of inattention and hyperactivity-impulsivity separately and across raters. ADHD is a heterogeneous disorder, and the two dimensions are not perfectly correlated at the phenotypic or genetic level (Greven *et al.* 2011*a*, b; Larsson *et al.* 2012*b*). Accordingly, one recent twin study found that parents and teachers rated unique aspects of inattentive and hyperactive-impulsive behaviors (McLoughlin *et al.* 2011).

There are two main implications that arise from this study. First, the identification of a highly heritable common factor suggests that clinical and etiological investigations of ADHD will benefit from combining data from multiple informants to create a pervasive, more heritable phenotype. This has the effect of reducing measurement error, thereby increasing power for tests of association with genetic, environmental and neurobiological variables. The second implication is for our understanding of the self-rating measures that are used in most adult studies of ADHD. Our findings suggest that self-ratings in childhood, when used as the sole measure of ADHD symptoms, may underestimate heritability. Thus, previous results indicating lower heritability of ADHD in adulthood may be due to a rater effect rather than a true change in the extent of genetic influences over the course of development. Longitudinal family and twin studies are now required to characterize stability and change in the familial and genetic influences on ADHD symptoms throughout the lifespan, and whenever possible should include similar informant ratings to those routinely collected in childhood and adolescence, in addition to self-ratings.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

References

Achenbach, TM.; Rescorla, LA. Manual for ASEBA School-Age Forms and Profiles. Department of Psychiatry, University of Vermont; Burlington, VT: 2000.

APA. Diagnostic and Statistical Manual of Mental Disorders. 4th edn. American Psychiatric Association; Washington, DC: 1994.

Asherson P. Clinical assessment and treatment of attention deficit hyperactivity disorder in adults. Expert Review of Neurotherapeutics. 2005; 5:525–539. [PubMed: 16026236]

Boomsma DI, Saviouk V, Hottenga JJ, Distel MA, de Moor MH, Vink JM, Geels LM, van Beek JH, Bartels M, de Geus EJ, Willemsen G. Genetic epidemiology of attention deficit hyperactivity disorder (ADHD index) in adults. PLoS One. 2010; 5:e10621. [PubMed: 20485550]

Burt SA. Rethinking environmental contributions to child and adolescent psychopathology: a meta-analysis of shared environmental influences. Psychological Bulletin. 2009; 135:608–637. [PubMed: 19586164]

Chen W, Zhou K, Sham P, Franke B, Kuntsi J, Campbell D, Fleischman K, Knight J, Andreou P, Arnold R, Altink M, Boer F, Boholst MJ, Buschgens C, Butler L, Christiansen H, Fliers E, Howe-Forbes R, Gabriëls I, Heise A, Korn-Lubetzki I, Marco R, Medad S, Minderaa R, Müller UC, Mulligan A, Psychogiou L, Rommelse N, Sethna V, Uebel H, McGuffin P, Plomin R, Banaschewski T, Buitelaar J, Ebstein R, Eisenberg J, Gill M, Manor I, Miranda A, Mulas F, Oades RD, Roeyers H, Rothenberger A, Sergeant J, Sonuga-Barke E, Steinhausen HC, Taylor E, Thompson M, Faraone SV, Asherson P. DSM-IV combined type ADHD shows familial association with sibling trait scores: a sampling strategy for QTL linkage. American Journal of Medical Genetics. Part B, Neuropsychiatric Genetics. 2008; 147:1450–1460.

Davis OSP, Arden R, Plomin R. g in middle childhood: moderate genetic and shared environmental influence using diverse measures of general cognitive ability at 7, 9 and 10 years in a large population sample of twins. Intelligence. 2008; 36:68–80.

Derks EM, Hudziak JJ, Van Beijsterveldt CE, Dolan CV, Boomsma DI. Genetic analyses of maternal and teacher ratings on attention problems in 7-year-old Dutch twins. Behavior Genetics. 2006; 36:833–844. [PubMed: 16773450]

Ehringer MA, Rhee SH, Young S, Corley R, Hewitt JK. Genetic and environmental contributions to common psychopathologies of childhood and adolescence: a study of twins and their siblings. Journal of Abnormal Child Psychology. 2006; 34:1–17. [PubMed: 16465480]

Faraone SV, Biederman J, Mick E. The age-dependent decline of attention deficit hyperactivity disorder: a meta-analysis of follow-up studies. Psychological Medicine. 2006; 36:159–165. [PubMed: 16420712]

Goodman R. Psychometric properties of the strengths and difficulties questionnaire. Journal of the American Academy of Child and Adolescent Psychiatry. 2001; 40:1337–1345. [PubMed: 11699809]

Greven CU, Asherson P, Rijsdijk FV, Plomin R. A longitudinal twin study on the association between inattentive and hyperactive-impulsive ADHD symptoms. Journal of Abnormal Child Psychology. 2011a; 39:623–632. [PubMed: 21494861]

Greven CU, Rijsdijk FV, Plomin R. A twin study of ADHD symptoms in early adolescence: hyperactivity-impulsivity and inattentiveness show substantial genetic overlap but also genetic specificity. Journal of Abnormal Child Psychology. 2011b; 39:265–275. [PubMed: 21336711]

Haberstick BC, Timberlake D, Hopfer CJ, Lessem JM, Ehringer MA, Hewitt JK. Genetic and environmental contributions to retrospectively reported DSM-IV childhood attention deficit hyperactivity disorder. Psychological Medicine. 2008; 38:1057–1066. [PubMed: 17892623]

Hartman CA, Rhee SH, Willcutt EG, Pennington BF. Modeling rater disagreement for ADHD: are parents or teachers biased? Journal of Abnormal Child Psychology. 2007; 35:536–542. [PubMed: 17333362]

Hewitt JK, Silberg JL, Neale MC, Eaves LJ, Erickson M. The analysis of parental ratings of children's behavior using LISREL. Behavior Genetics. 1992; 22:293–317. [PubMed: 1616461]

Lamb DJ, Middeldorp CM, Van Beijsterveldt CE, Boomsma DI. Gene–environment interaction in teacher-rated internalizing and externalizing problem behavior in 7- to 12-year-old twins. Journal of Child Psychology and Psychiatry. 2012; 53:818–825. [PubMed: 22519827]

Larsson H, Anckarsäter H, Råstam M, Chang Z, Lichtenstein P. Childhood attention-deficit hyperactivity disorder as an extreme of a continuous trait: a quantitative genetic study of 8,500 twin pairs. Journal of Child Psychology and Psychiatry. 2012a; 53:73–80. [PubMed: 21923806]

Larsson H, Asherson P, Chang Z, Ljung T, Friedrichs B, Larsson JO, Lichtenstein P. Genetic and environmental influences on adult attention deficit hyperactivity disorder symptoms: a large Swedish population-based study of twins. Psychological Medicine. 2012b Published online: 6 August 2012. doi:10.1017/S0033291712001067.

Loehlin JC. The Cholesky approach: a cautionary note. Behavior Genetics. 1996; 26:65–69.

Martin N, Scourfield J, McGuffin P. Observer effects and heritability of childhood attention-deficit hyperactivity disorder symptoms. British Journal of Psychiatry. 2002; 180:260–265. [PubMed: 11872519]

McGue M, Bouchard TJ Jr. Adjustment of twin data for the effects of age and sex. Behavior Genetics. 1984; 14:325–343. [PubMed: 6542356]

McLoughlin G, Rijsdijk F, Asherson P, Kuntsi J. Parents and teachers make different contributions to a shared perspective on hyperactive-impulsive and inattentive symptoms: a multivariate analysis of parent and teacher ratings on the symptom domains of ADHD. Behavior Genetics. 2011; 41:668–669. [PubMed: 21660593]

Nadder TS, Rutter M, Silberg JL, Maes HH, Eaves LJ. Genetic effects on the variation and covariation of attention deficit-hyperactivity disorder (ADHD) and oppositional-defiant disorder/conduct disorder (ODD/CD) symptomatologies across informant and occasion of measurement. Psychological Medicine. 2002; 32:39–53. [PubMed: 11883729]

Neale, MC. Mx: Statistical Modeling. 4th edn. Department of Psychiatry, Medical College of Virginia; Richmond, VA: 1997.

Neale MC, Miller MB. The use of likelihood-based confidence intervals in genetic models. Behavior Genetics. 1997; 27:113–120. [PubMed: 9145549]

Nikolas MA, Burt SA. Genetic and environmental influences on ADHD symptom dimensions of inattention and hyperactivity: a meta-analysis. Journal of Abnormal Psychology. 2010; 119:1–17. [PubMed: 20141238]

Oliver BR, Plomin R. Twins' Early Development Study (TEDS): a multivariate, longitudinal genetic investigation of language, cognition and behavior problems from childhood through adolescence. Twin Research and Human Genetics. 2007; 10:96–105. [PubMed: 17539369]

Plomin, R.; DeFries, JC.; McClearn, GE.; McGuffin, P. Behavioral Genetics. Worth Publishers; New York: 2008.

Price TS, Freeman B, Craig I, Petrill SA, Ebersole L, Plomin R. Infant zygosity can be assigned by parental report questionnaire data. Twin Research. 2000; 3:129–133. [PubMed: 11035484]

Price TS, Simonoff E, Asherson P, Curran S, Kuntsi J, Waldman I, Plomin R. Continuity and change in preschool ADHD symptoms: longitudinal genetic analysis with contrast effects. Behavior Genetics. 2005; 35:121–132. [PubMed: 15685426]

Price TS, Simonoff E, Waldman I, Asherson P, Plomin R. Hyperactivity in preschool children is highly heritable. Journal of the American Academy of Child and Adolescent Psychiatry. 2001; 40:1362–1364. [PubMed: 11765277]

Raftery AE. Bayesian model selection in social research. Sociological Methodology. 1995; 25:111–163.

Rietveld MJH, Posthuma D, Dolan CV, Boomsma DI. ADHD: sibling interaction or dominance: an evaluation of statistical power. Behavior Genetics. 2003; 33:247–255. [PubMed: 12837015]

Saudino KJ, Ronald A, Plomin R. The etiology of behavior problems in 7-year-old twins: substantial genetic influence and negligible shared environmental influence for parent ratings and ratings by same and different teachers. Journal of Abnormal Child Psychology. 2005; 33:113–130. [PubMed: 15759595]

Saviouk V, Hottenga JJ, Slagboom EP, Distel MA, de Geus EJ, Willemsen G, Boomsma DI. ADHD in Dutch adults: heritability and linkage study. American Journal of Medical Genetics. Part B, Neuropsychiatric Genetics. 2011; 156B:352–362.

Schultz MR, Rabi K, Faraone SV, Kremen W, Lyons MJ. Efficacy of retrospective recall of attention-deficit hyperactivity disorder symptoms: a twin study. Twin Research and Human Genetics. 2006; 9:220–232. [PubMed: 16611492]

Simonoff E, Pickles A, Hervas A, Silberg JL, Rutter M, Eaves L. Genetic influences on childhood hyperactivity: contrast effects imply parental rating bias, not sibling interaction. Psychological Medicine. 1998; 28:825–837. [PubMed: 9723139]

StataCorp. Stata Statistical Software: Release 10. StataCorp LP; College Station, TX: 2007.

Thapar A, Harrington R, Ross K, McGuffin P. Does the definition of ADHD affect heritability? Journal of the American Academy of Child and Adolescent Psychiatry. 2000; 39:1528–1536. [PubMed: 11128330]

Van Den Berg SM, Willemsen G, De Geus EJC, Boomsma DI. Genetic etiology of stability of attention problems in young adulthood. American Journal of Medical Genetics. Part B, Neuropsychiatric Genetics. 2006; 141B:55–60.

Van Roy B, Veenstra M, Clench-Aas J. Construct validity of the five-factor Strengths and Difficulties Questionnaire (SDQ) in pre-, early, and late adolescence. Journal of Child Psychology and Psychiatry. 2008; 49:1304–1312. [PubMed: 19120709]

Williams RL. A note on robust variance estimation for cluster-correlated data. Biometrics. 2000; 56:645–646. [PubMed: 10877330]

Wood AC, Buitelaar J, Rijsdijk F, Asherson P, Kuntsi J. Rethinking shared environment as a source of variance underlying attention-deficit/hyperactivity disorder symptoms: comment on Burt (2009). Psychological Bulletin. 2010; 136:331–340. [PubMed: 20438137]

Young SE, Stallings MC, Corley RP, Krauter KS, Hewitt JK. Genetic and environmental influences on behavioral disinhibition. American Journal of Medical Genetics. Part B, Neuropsychiatric Genetics. 2000; 96:684–695.
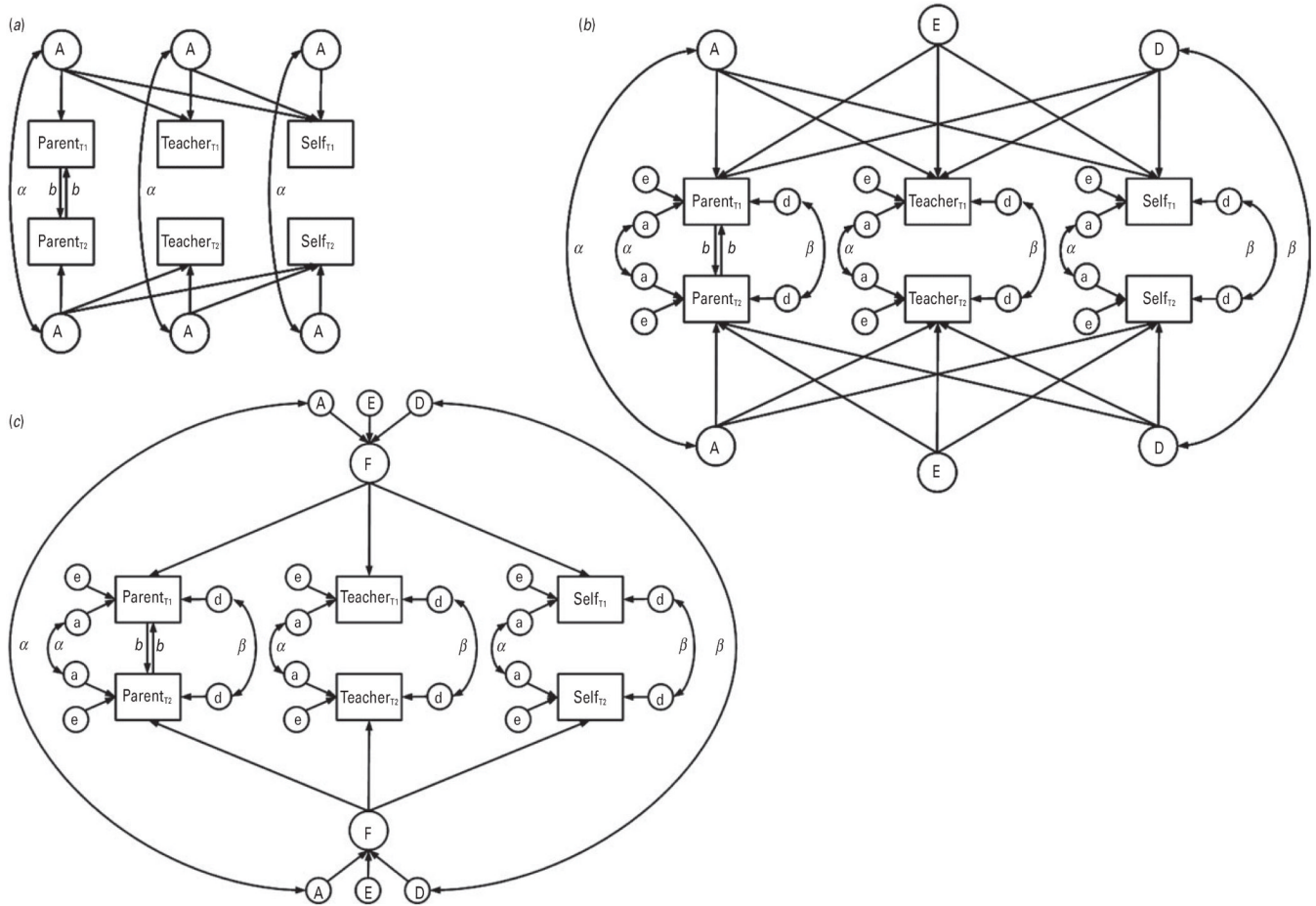
Fig. 1.
Path diagrams for the multivariate genetic models. (*a*) Correlated factors solution: depicts only the additive genetic (A) correlations between parent, teacher and self-ratings of attention deficit hyperactivity disorder (ADHD), for illustrative purposes; non-additive genetic (D) and non-shared environmental (E) correlations are similar, except that the coefficient of relatedness between twins is for D and 0 for E. (*b*) Independent pathway model: A, D and E are common components of variance; a, d and e are residual components of variance unique to each informant. (*c*) Common pathway model: A, D and E are components of variance influencing the common latent factor (*F*); a, d and e are residual components of variance unique to each informant. T_1, twin 1; T_2, twin 2; A, additive genetic component of variance; *b*, contrast effect for parent ratings; , coefficient of additive genetic relatedness between T_1 and T_2, set to 1.00 for monozygotic (MZ) pairs and 0.5 for dizygotic (DZ) pairs; , coefficient of non-additive genetic relatedness between T_1 and T_2, set to 1.00 for MZ pairs and 0.25 for DZ pairs.

Table 1

Descriptive statistics and correlations

	MZM	MZF	DZM	DZF	DZO
n pairs					
P	908	1116	841	976	1749
T	862	1014	781	923	1637
S	918	1113	845	982	1763
Mean (s.d.)					
P	3.36 (2.25)	2.29 (1.96)	3.23 (2.39)	2.50 (2.14)	2.81 (2.31)
T	2.98 (2.74)	1.48 (1.90)	2.92 (2.76)	1.66 (2.06)	2.20 (2.53)
S	3.81 (2.37)	3.10 (2.12)	3.89 (2.37)	3.29 (2.24)	3.58 (2.31)
Variance					
P	0.30 (0.28–0.31)	0.29 (0.28–0.29)	0.33 (0.31–0.36)	0.32 (0.30–0.34)	0.35 (0.33–0.37)/0.28 (0.26–0.30)
T	0.46 (0.42–0.49)	0.30 (0.28–0.32)	0.48 (0.44–0.51)	0.33 (0.31–0.35)	0.48 (0.45–0.52)/0.31 (0.29–0.33)
S	0.32 (0.30–0.34)	0.30 (0.28–0.32)	0.30 (0.29–0.32)	0.32 (0.30–0.34)	0.31 (0.29–0.33)/0.29 (0.27–0.31)
rWT					
P	0.75 (0.72–0.78)	0.77 (0.74–0.79)	0.23 (0.17–0.29)	0.32 (0.26–0.32)	0.25 (0.21–0.29)
T	0.63 (0.58–0.67)	0.57 (0.53–0.57)	0.29 (0.22–0.36)	0.33 (0.27–0.39)	0.31 (0.26–0.35)
S	0.49 (0.44–0.53)	0.48 (0.44–0.52)	0.21 (0.15–0.27)	0.21 (0.15–0.27)	0.15 (0.11–0.19)
rCTCT					
P and T	0.29 (0.28–0.32)	0.30 (0.28–0.33)	0.08 (0.03–0.14)	0.12 (–0.08 to 0.17)	0.09 (0.06–0.12)
P and S	0.35 (0.32–0.38)	0.37 (0.35–0.40)	0.07 (0.02–0.12)	0.12 (0.08–0.17)	0.09 (0.06–0.12)
T and S	0.26 (0.23–0.30)	0.25 (0.21–0.28)	0.09 (0.04–0.14)	0.16 (0.15–0.20)	0.10 (0.07–0.14)

MZM, monozygotic male; MZF, monozygotic female; DZM, dizygotic male; DZF, dizygotic female; DZO, dizygotic opposite-sex; P, parent; T, teacher; S, self; *r*WT, within-twin correlation; *r*CTCT, cross-twin cross-trait correlation.

n includes complete and incomplete pairs; mean and standard deviation (s.d.) reported for raw (untransformed) data, variances and correlations reported for transformed data regressed on age and sex; variances reported separately for male and female DZO twins; 95 % confidence intervals for variances and correlations in parentheses.

Table 2

Parameter estimates for the common pathway model (whole sample)

	Latent factor	Parent	Teacher	Self
F	–	0.72 (0.69–0.75)	0.46 (0.44–0.49)	0.63 (0.60–0.66)
A^2	0.34 (0.13–0.56)	–	–	–
D^2	0.49 (0.28–0.71)	–	–	–
E^2	0.16 (0.14–0.19)	–	–	–
a^2	–	0.36 (0.32–0.39)	0.43 (0.43–0.47)	0.16 (0.12–0.19)
e^2	–	0.12 (0.10–0.14)	0.36 (0.36–0.49)	0.45 (0.42–0.45)
b	–	−0.04 (−0.06 to −0.02)	–	–

Proportion of phenotypic variance (%)

	Parent	Teacher	Self
Common A	18	7	13
Common D	25	10	19
Common E	8	3	6
Residual a	36	43	16
Residual e	12	36	45

F is the loading of each informant rating onto the latent factor; A^2, D^2 and E^2 are standardized components of variance for latent factor; a^2 and e^2 are standardized components of variance unique to each informant rating; b is the rater contrast effect unique to parent ratings; the lower section of the table gives the proportion of phenotypic variance explained by common/residual genetic (A/a, D) and non-shared environmental (E/e) factors for each informant, where the proportion of variance from the common factor is calculated as the factor loading multiplied by the standardized parameter estimate multiplied by the factor loading (i.e. Common A = F × A^2 × F); 95 % confidence intervals in parentheses.

Table 3

Parameter estimates for the same-teacher and different-teacher common pathway models

	Same teacher				Different teacher			
	Latent factor	Parent	Teacher	Self	Latent factor	Parent	Teacher	Self
F	–	0.74 (0.70–0.79)	0.45 (0.41–0.49)	0.61 (0.57–0.66)	–	0.71 (0.68–0.75)	0.47 (0.44–0.50)	0.63 (0.60–0.67)
A^2	0.28 (0.00–0.57)	–	–	–	0.34 (0.07–0.57)	–	–	–
D^2	0.57 (0.28–0.87)	–	–	–	0.49 (0.22–0.76)	–	–	–
E^2	0.15 (0.11–0.20)	–	–	–	0.17 (0.14–0.21)	–	–	–
a^2	–	0.30 (0.23–0.36)	0.58 (0.54–0.62)	0.18 (0.13–0.24)	–	0.39 (0.34–0.44)	0.31 (0.26–0.35)	0.14 (0.10–0.19)
e^2	–	0.15 (0.12–0.18)	0.22 (0.19–0.25)	0.44 (0.40–0.48)	–	0.10 (0.08–0.13)	0.47 (0.43–0.51)	0.45 (0.42–0.49)
b	–	–	–	–	–	–0.05 (–0.08 to –0.04)	–	–

Proportion of phenotypic variance (%)

	Parent	Teacher	Self
Common A	15	6	10
Common D	31	12	21
Common E	8	3	6
Residual a	30	58	18
Residual e	15	22	44

Proportion of phenotypic variance (%)

	Parent	Teacher	Self
Common A	17	8	13
Common D	25	11	19
Common E	9	4	7
Residual a	39	31	14
Residual e	10	47	45

F is the loading of each informant rating onto the latent factor; A^2, D^2 and E^2 are standardized components of variance for latent factor; a^2 and e^2 are standardized components of variance unique to each informant rating; b is the rater contrast effect unique to parent ratings; the lower section of the table gives the proportion of phenotypic variance explained by common/residual genetic (A/a, D) and non-shared environmental (E/e) factors for each informant, where the proportion of variance from the common factor is calculated as the factor loading multiplied by the standardized parameter estimate multiplied by the factor loading (i.e. Common A = F × A^2 × F); 95 % confidence intervals in parentheses.