Efficacy of Retrospective Recall of Attention-Deficit Hyperactivity Disorder Symptoms: A Twin Study

Mark R. Schultz,[1] Keren Rabi,[1] Stephen V. Faraone,[2] William Kremen,[3] and Michael J. Lyons[1]

[1] *Psychology Department, Boston University, Boston, United States of America*
[2] *Medical Genetics Research Center, Department of Psychiatry, SUNY Upstate Medical University, Syracuse, United States of America*
[3] *Department of Psychiatry, University of California, San Deigo, La Jolla, California, United States of America*

Attention-deficit hyperactivity disorder (ADHD) is currently recognized as a neurobiological, genetically based disorder in both children and adults. In this article we examine whether, by using a sample of middle-aged male twin veterans, the phenotypic characterization, prevalence, heritability and the longitudinal course of the illness is comparable to results observed in samples of children and adolescents. We evaluated the utility of adult reports of lifetime ADHD symptoms by examining the heritability of retrospectively reported childhood symptoms, using both symptom-based and discrete classification-based approaches, as well as examining the persistence of ADHD symptoms into adulthood for that subsample of individuals who were judged to possibly have ADHD as children. Our results showed prevalence rates that were approximately similar to those observed in other studies, demonstrable familiality, similar item endorsement patterns, a strong genetic association between hyperactive and inattentive subtypes, and a longitudinal decline in symptom severity. We concluded that while assessing ADHD in adult probands may be less accurate than with children or adolescents, since it demonstrates several characteristics in common with other assessment techniques it remains a viable diagnostic and research strategy, even with population samples.

Attention-deficit hyperactivity disorder (ADHD) is a neurobiological (Faraone, 2004a, 2004b; Seidman et al., 1998), genetically influenced (e.g., Faraone et al., 2005; Todd, 2000) disorder. Empirical studies have demonstrated a pattern of consistent decline in ADHD symptoms as the individual patient ages (Biederman et al., 1996). While estimates of prevalence vary, typical figures are 8% to 12% among children and adolescents (Faraone et al., 2003), and somewhat less in adults (Barkley, 1998; Kessler et al., 1999). Key clinical features of the syndrome include various forms of inattentiveness and inappropriate motor activity. Typically, as the individual ages, the severity and number of symptoms he or she displays decrease.

Genetic studies have shown ADHD to be among the most heritable of childhood psychiatric disorders (Faraone et al., 2005; Todd, 2000). Twin studies are recognized as the 'gold standard' for assessing the degree to which ADHD is influenced by genetic factors (Faraone, 2004a; Gillis et al., 1992; Sharp et al., 2003; Thapar et al., 1999) as they provide a reliable method of assessing the heritability of ADHD (Faraone, 2004b). Such studies consistently indicate ADHD symptoms to be highly heritable, typically demonstrating 70% to 80% heritability (Eaves et al., 1997; Faraone & Biederman, 1994; Martin et al., 2002; Thapar et al., 2000). However, there are notable exceptions to this estimate (Sherman, Iacono, et al., 1997). Sherman, McGue, et al. (1997) found a broad range of heritability estimates, from a low of .39 (teacher ratings of inattention) to a high of .91 (mother ratings of hyperactivity).

Some researchers have suggested that the consistency of results from genetic studies depends on a uniform phenotype definition as well as on the type of rater used (e.g., Martin et al., 2002). Nadder et al. (2001), for example, showed a rater-specific variance that influenced genetic variance estimates. They maintained that twin studies should obtain reports from multiple informants to avoid rater bias effects that may substantially influence heritability estimates (Sherman, Iacono, et al., 1997). Martin et al. (2002) conducted an adolescent twin study using self- and multiple informants' ratings. Self-reported symptom data showed no evidence of genetic effects. However, when ADHD was assessed by parents and teachers, a high heritability for ADHD was demonstrated. Other studies (Goodman & Stevenson, 1989; Sherman,

Received 31 August, 2005; accepted 3 November, 2005.

Address for correspondence: Mark Schultz, Psychology Department, Boston University, 648 Beacon Street, Boston, MA 02215, USA. E-mail: mschult@bu.edu

McGue, et al., 1997) demonstrated that parents seem to perceive monozygotic (MZ) twins as far more similar than teachers do, suggesting that the poor fit of the resulting additive genetic models (MZ correlations >> dizygotic [DZ]) may indicate substantial rater bias (on the part of the parents). Additional twin studies have demonstrated that rater effects may bias genetic estimates (Eaves et al., 1997; Martin et al., 2002; Thapar et al., 2000). Sherman, McGue, et al. (1997) found that differences between raters resulted in high genetic correlations for mother's reports (.86) between hyperactivity and inattention subtypes but substantially less so (.33) for teacher's ratings.

The *Diagnostic and Statistical Manual of Mental Disorders* (3rd ed., rev.; DSM-III-R; American Psychiatric Association, 1987) conceptualized ADHD as a unidimensional syndrome. However, more recent work (American Psychiatric Association, 1994; Todd, 2000) has moved away from that perspective, typically proposing a bidimensional model consisting of strongly related inattentive and hyperactive symptoms. For example, Sherman, McGue, et al. (1997) found moderate phenotypic correlations (.57 for mothers and .68 for teachers) between the dimensions of hyperactivity and inattention. Bauermeister et al. (1992) also found support for a bidimensional conceptualization of ADHD using a combination of factor analytic and cluster analytic techniques. A bidimensional view of the syndrome is consistent with Rassmussen et al.'s (2002) assertion that the technique of latent class analysis produces distinct clinically relevant classes of inattention, hyperactive-impulsive, and combined type, similar to the classifications in the fourth edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-IV; American Psychiatric Association, 1994) classifications. Moreover, some researchers (Neuman et al., 1999; Todd, 2000) have suggested that the qualitative, categorical diagnosis of the syndrome is problematic, suggesting instead the use of latent class categories or alternative (i.e., continuous) phenotype definitions of the disorder.

Ample evidence suggests that ADHD is a chronic condition in many cases (Biederman et al., 1996; Faraone et al., 2000; Mannuzza et al., 1993). Follow-up studies indicate a substantial ADHD symptomatic persistence into adulthood (Faraone et al., in press), with some studies suggesting that as many as 80% to 90% of children with ADHD continue to experience impairing symptoms of the disorder into young adulthood (Barkley, 1998; Biederman et al., 2000) even though empirical studies demonstrate a pattern of decline in ADHD symptoms (particularly hyperactivity-impulsive symptoms) with age (Biederman et al., 1996; Biederman et al., 2000; Faraone et al., in press; Mannuzza et al., 1993).

Although adult informants are typically used to validate the symptoms of children and adolescents when assessing ADHD, they are not typically available when assessing ADHD in adults. As a result, the clinical diagnosis of adult ADHD commonly relies on the patient's account of retrospective (childhood) and current symptoms. Since current diagnostic criteria require an early onset of ADHD symptoms (American Psychiatric Association, 1994), diagnosis critically depends on the adult's ability to provide accurate and detailed memories of long-ago events and behavior (Mannuzza et al., 2002). Consequently, the reliability and validity of such retrospective recalls are of major concern in epidemiological and etiological research (Mannuzza et al., 2002).

The potential limitations of retrospective self-report of childhood ADHD symptoms during early adulthood have been extensively examined. Mannuzza et al. (2002) conducted a longitudinal follow-up study of adolescents who had received an ADHD diagnosis during childhood with the intent of examining the accuracy of their retrospective recalls of childhood ADHD symptoms at the age of 25 years. Retrospective ADHD diagnoses yielded a sensitivity and specificity of .78 and .89, respectively. These authors concluded that evaluating adult ADHD using retrospective self-reports would most likely result in valid diagnostic results. The authors cautioned that when retrospectively assessing childhood ADHD in population (as opposed to clinical) samples of adults, such as in epidemiological surveys, estimates of prevalence would most likely be overestimated (Mannuzza et al., 2002) since many of the diagnosed cases would be false positives due to the low prevalence of the disorder.

Barkley et al. (2002) demonstrated substantially higher ratings of adult ADHD symptoms when using parent reports compared to using the proband's symptomatic recollections. The parent reports were seen as more valid as they were better predictors of current impairment than the self-reports. The authors suggested that previous follow-up studies that relied on adult self-reports might have substantially underestimated the persistence of ADHD into adulthood.

Zucker et al. (2002) used a sample of clinically referred college students with ADHD to evaluate the concordance of self- and informant ratings of current and childhood ADHD symptoms. Their results demonstrated a moderate concordance for both current and childhood symptoms, with informants consistently scoring the proband higher for inattentive symptoms (hyperactivity-impulsivity were scored similarly by both self- and informant ratings). In addition, the study also suggested that inattentive symptoms are the dominant syndrome features during adulthood (Zucker et al., 2002). Conversely, Murphy and Schachar (2000) suggested that the diagnosis of adult ADHD can be validly established even when solely based on the retrospective recall of childhood behavioral symptoms since they found no mean differences in the current or past symptomatic ratings of nonclinical adult samples when compared to their informants' ratings.

Most studies of the syndrome have used third party informants for children (parents or teachers) or

https://doi.org/10.1375/twin.9.2.220 Published online by Cambridge University Press

self-assessment with older individuals. Unlike most of the work done on ADHD, we studied a middle-aged cohort (the Vietnam Era Twin Registry) to demonstrate the viability of using retrospective self-report. In this study we replicate and extend the work of Murphy and Schachar (2000), evaluating the validity of using adult reports of lifetime ADHD symptoms. We examined the prevalence and heritability of retrospectively reported childhood symptoms as well as the persistence of these symptoms into midlife. We concluded that the validity of retrospective self-reporting would be supported if the following four criteria were met: (1) the heritability estimates observed from adult self-report data are consistent with other studies that have used young probands or third party informants, (2) the pattern of item endorsement using retrospective reporting is similar to other studies that used more traditional assessment, (3) the relationship between subtypes (inattentive, hyperactive, combined) with a retrospectively assessed adult sample replicates the relationship among types found in juvenile samples, and (4) since most prospective studies of ADHD show a substantial decline in symptoms as the proband ages, a similar pattern of results in our self-reported data would enhance the probability that our subjects are reporting their symptoms accurately.

Materials and Methods

Participants were drawn from the Vietnam Era Twin Registry, a national registry of over 10,000 male twins (Eisen et al., 1987) in which both members served in the military between 1965 to 1975. The present sample included 345 twin pairs and 2 singletons (692 individuals) between the ages of 41 and 58 recruited from more than 3300 pairs surveyed in a previous study of drug abuse and dependence (Tsuang et al., 2001). Participants in this study were randomly selected from the pool of participants in the previous study with two additional restrictions: (1) both members of a pair agree to participate, and (2) neither member of a twin pair has a history of exposure to combat. Participants were selected independently of substance use disorders or other psychiatric or medical conditions.

The mean age of all participants was 47.8 years (*SD* = 3.3); 92.2% were Caucasian, 5.5% were African American, 1.9% were Hispanic, and 0.4% were of other racial origin. In addition, 96.7% were high school graduates and 33% were college graduates; 79.1% were married, 12.1% divorced, and 8.8% widowed, separated, never married, or refused response. Among participants reporting full-time (92.2%) or part-time (1.6%) employment, 33.5% held service or manual labor positions, 24.4% held clerical or semiprofessional positions, and 41.1% held professional positions. Median household income category was $60,000 to $70,000 (range, <$10,000–$100,000+). Zygosity was determined using questionnaire and blood group methods with 95% accuracy (Eisen et al., 1989). A complete description of the registry's construction has been previously reported (Eisen et al., 1987).

The subjects were administered a structured interview about ADHD symptoms (Schedule for Affective Disorders and Schizophrenia for School Age Children; K-SADS-E; Orvaschel & Puig-Antich, 1987). Subjects were asked to recall the applicability of each of 18 symptoms covered by the K-SADS-E questionnaire based on their behavior as children on a 4-point scale (1 = *none*, 2 = *mild*, 3 = *moderate*, 4 = *severe*). If the subject endorsed two or more symptoms (approximately 20% did so) to a clinically meaningful degree (at least two endorsements of *mild* impairment), they were asked whether the same symptoms were currently present.

Seventeen of the 692 subject interviews failed to provide complete data for all items of the K-SADS-E. Of these, most failed to respond to only one or two questions, with less than 0.5% of the interviews containing four or more omissions. To avoid loss of data, we recoded missing data conservatively (for example, if a subject failed to respond to an item, that item was coded at its lowest — 'subject does not endorse symptom' — response value). This allowed us to include 13 subjects' data that would otherwise have been discarded.

The K-SADS-E items are a superset of the items that compose the DSM-III-R ADHD criteria (14 items) but, unlike the DSM-III-R, the K-SADS-E is comprised of three subscales (hyperactivity, inattention, and other). Four K-SADS-E items (has lots of accidents, clumsy, fights with peers, rejected by peers) did not correspond to the DSM-III-R items and were consequently excluded from analyses.

The DSM-III-R criteria for the diagnosis of ADHD requires an endorsement of eight or more symptoms (at a *moderate* or higher level of severity) out of 14 possible. Applying those criteria to children, prevalence rates of 8% to 12% have typically been reported (Faraone et al., 2003), while adult rates of 5% have been reported in other samples (Heiligenstein et al., 1998; Kessler et al., in press; Murphy & Barkley, 1996). To increase our statistical power, we followed a procedure proposed by Levy et al. (1997), who examined the effect of defining 'broad diagnostic criteria' based on a subject endorsing five or more symptoms rather than the eight called for by DSM-III-R. Levy et al. (1997) validated the use of the five-symptom threshold by comparing both thresholds against ratings produced using the Diagnostic Interview Schedule for Children-Parent version (PC-DISC; Emory University, 1993). Using the classifications produced by the PC-DISC as the validity criterion, the 5+ symptom threshold produced a sensitivity and specificity of 79.4% and 87%, respectively whereas the 8+ threshold produced a sensitivity and specificity of 50% and 98.7%, respectively. Thus, we believed an endorsement of five or more symptoms at a *moderate* level represented a meaningful threshold for

https://doi.org/10.1375/twin.9.2.220 Published online by Cambridge University Press

ADHD, defined as a lifetime history of ADHD with onset in childhood.

We examined both the DSM-III-R subset of the K-SADS-E and the K-SADS-E subscales for inattention and hyperactivity/impulsivity due to questions regarding the multidimensionality of ADHD symptom counts.

Statistical Analyses

Concordance and Correlations

The basic approach of the twin method is to compare the degree of similarity within MZ twin pairs to the degree of similarity within DZ twin pairs. One method for quantifying similarity for dichotomous characteristics is the concordance rate. A finding of significant difference in concordance rates, with the MZ rate higher than the DZ, is evidence for a genetic influence. There are two commonly used approaches for calculating concordance rates, the proband method and the pairwise method. The pairwise concordance rate is calculated by dividing the number of concordant pairs by the number of pairs in which one or both of the twins has the disorder. The probandwise concordance rate is calculated by dividing twice the number of concordantly ill pairs by the number of pairs in which one or both twins has the phenotype of interest plus the number of concordantly ill pairs. The probandwise method provides rates that are comparable to risk rates from other types of family pairings or population prevalence figures (McGue, 1992). Probandwise concordance rates were calculated for the dichotomous phenotypes of DSM-III-R childhood ADHD using the 5+ symptom threshold. In addition, concordance rates for the inattentive and hyperactive subscales of the K-SADS-E were also computed. Since no thresholds similar to DSM-III-R existed to dichotomize these subscales into 'affected' vs. 'unaffected', we created our own liability thresholds for the individual subscale symptom count data that resulted in approximately the same relative distribution of 'affected vs. nonaffected' as the 5+ symptom endorsement of the DSM scale.

The concordance rate does not use all available information as only pairs in which at least one member is affected are included in its calculation. However, the tetrachoric correlation (Chen & Popovitch, 1995), calculated from the 2 x 2 contingency table for dichotomous outcomes, is an alternative statistic which uses all available information about the resemblance within twin pairs. Pairs that are concordant for being unaffected as well as affected are included in its computation. The tetrachoric correlation is also known as the correlation of liability (Falconer, 1965).

Model Fitting

Biometric modeling analyses were performed to test whether genetic effects or family environmental effects alone are responsible for observed familial similarity. Since the DSM classification (affected/not affected) does not use all of the information contained in the data, and since Neale et al. (1994) pointed out that the sample size requirements of the categorical approach to maintain equivalent power are substantially larger than for the continuous model, we also examined the retrospectively assessed childhood hyperactive and inattentive symptom counts instead of the DSM-III-R dichotomous diagnoses. However, fitting biometric models to psychiatric symptom scores can also be problematic for a number of reasons (van den Oord et al., 2000; van den Oord & van der Ark, 1997). For example, since in a population sample the vast majority of subjects score near the 'healthy' end of the scale, the twin correlations become attenuated due to censoring or 'floor' effects. These authors emphasized the need to deal with the type of censoring that occurs for sums of Likert items (which would be particularly severe for the DSM) because the censoring may seriously bias parameter estimates and distort goodness-of-fit tests. This can be problematic as van den Oord's simulation modeling (2000) demonstrated that attenuation of the correlations between symptom counts forced a general underestimation of genetic variance components and a general overestimation of unique environmental (and error) variance components. Neale et al. (1994) and Derks et al. (2004) suggested that by simultaneously estimating the thresholds and parameter estimates, most of these biases can be eliminated.

The statistical software program Mx (Neale et al., 2002) permits the use of raw ordinal data and multiple category thresholds to recover much of the power lost with the dichotomous approach. We fit univariate biometric models that estimated response thresholds to the symptom count data for each of the variables (DSM-III-R, hyperactive, and inattentive). For comparison, a categorical analysis was also conducted using the dichotomous disease status model for the retrospective childhood assessments of ADHD symptoms.

These models explain the observed phenotypic twin similarity in terms of additive genetic effects (heritability or h^2), common, shared, or family environmental effects (c^2), and unique or nonshared environmental effects (e^2). Parameters are estimated by the method of maximum likelihood using the Mx software package (Neale et al., 2002).

When conducting model-fitting to raw data in Mx, the χ^2 goodness-of-fit statistic is calculated by subtracting the log-likelihood of the fitted model from the log-likelihood of the observed data under a saturated model. The reproductive property of the χ^2 distribution (Dobson, 2002) ensures that the difference will also be distributed as a χ^2. When models are nested, the difference in fit between models can be tested by the difference in the χ^2 values ($\Delta\chi^2$) using as its degrees of freedom the *df* difference of the two models. If the change in χ^2 is not statistically significant, the more parsimonious model may be selected, as the test indicates that the constrained model fits equally well to the data.

https://doi.org/10.1375/twin.9.2.220 Published online by Cambridge University Press

Table 1

Three Measures of Association Computed Separately for MZ and DZ Twins (Spearman rho Correlations on Symptom Counts, Concordances and Tetrachoric Correlations) for Three Retrospective Childhood Phenotypic Assessments of ADHD; A 'Broad' DSM-III-R Criterion, and the Hyperactivity and Inattention Subscales of the K-SADS-E

		Spearman rho r_{xy} (symptom counts)	Tetrachoric corr. (dichotomous)	Concordance rate (dichotomous)
DSM-III-R	Monozygotic	.33	.54	.29
	Dizygotic	.17	.40	.23
Hyperactivity	Monozygotic	.28	.57	.37
	Dizygotic	.17	.40	.22
Inattention	Monozygotic	.34	.55	.35
	Dizygotic	.17	.48	.28

Note: For the correlation calculations, the $N = 169$ for DZ twins and $N = 173$ for MZ.

ADHD is currently viewed as a multidimensional disorder (American Psychiatric Association, 1994). Thus, we employed bivariate biometric analysis (again using raw data and multiple thresholds) to examine the covariance between the traits of hyperactivity and inattention as measured by the K-SADS-E subscales. These analyses allowed us to assess the genetic and environmental contributions to the K-SADS-E subscale data individually as well as examining the genetic and environmental contributions to the correlation between them.

Comparison of Endorsement Rates

Thissen et al. (1995) defined 'item difficulty' as inversely proportional to the percentage of subjects who endorse the item. Thus an item that is endorsed by 5% of subjects is substantially more 'difficult' than one endorsed by 30%. Even though an item with a low difficulty will be endorsed by a greater percent of affected subjects, it may not discriminate well between those that are affected and not. Mannuzza et al. (2002) examined the relative diagnostic utility of individual symptoms by applying the following criteria: (1) high odds ratio (greater than 10.00) and (2) sensitivity and specificity both greater than .70. They found six symptoms of ADHD that fulfilled both criteria (distractibility, concentration difficulties, complaints of inattention, acting before thinking, being 'on the go,' and fidgeting and squirming). We examined the endorsement rates of the individual K-SADS-E items for the subset of subjects that were given the 'current' interview and compared them to Mannuzza's result. In addition, we created an ad hoc scale from these 'frequently endorsed' items and computed its heritability.

'Adult' Versus 'Child' Evaluations

Those subjects who endorsed any childhood symptoms to a clinically meaningful extent (which we defined operationally as at least two symptoms rated *mild* or greater) were judged to display mild levels of ADHD symptoms and were given the same structured interview again and asked to rate the applicability of the symptoms to themselves as adults.

Selecting only those subjects who had a DSM-III-R retrospective childhood rating of at least five symptoms at a moderately severe level or greater (the Levy et al., 1997, criterion for 'affected'), we compared the current (adult) ratings of this subset of subjects deemed to be affected as children using paired sample *t* tests to determine whether changes in the number or severity of self-assessed ADHD symptoms might have occurred. Nonparametric correlations were also computed (rank ordering of self-assessed childhood rating vs. rank ordering of self-assessed adult rating) to examine the symptom stability over time.

Results

DSM-III-R specifies that a diagnosis of ADHD requires the presence of at least 8 of 14 diagnostic criteria with onset occurring before age 7. When retrospectively assessed, this definition produced rather low prevalence rates; 2.3% of our sample met criteria for retrospective reports of ADHD in childhood. The adult prevalence (same questions with emphasis on current *adult* behavior) in our sample was 0.5%. We then applied Levy's 'broad diagnostic category' criteria for diagnosing childhood ADHD (subject endorsed five or more items instead of eight; Levy et al., 1997). A McNemar test for association was nonsignificant, which showed the prevalence of ADHD between A and B twins (designations randomly assigned) did not differ. Using Levy's 'broad diagnostic criteria' the prevalence of childhood onset ADHD (irrespective of continuance of symptoms to adulthood) in our sample was 6.9%. This lower threshold yielded probandwise concordance rates of .29 and .23 for the MZ and DZ twins on the DSM-III-R measure, respectively, as shown in Table 1.

Prevalence rates were also calculated for the inattentive and hyperactive subtypes of ADHD as defined by the K-SADS-E criteria. Since the DSM-III-R does not specify symptom cutoff criteria for subtypes, we imposed the constraint on the K-SADS-E symptom counts that the 'affected' threshold be an integer number of symptoms and yield a roughly comparable

https://doi.org/10.1375/twin.9.2.220 Published online by Cambridge University Press

Table 2

Results of Biometric Modeling on Retrospectively Assessed Childhood Symptom Counts for the Three Scales of Symptom Counts: DSM-III-R, Hyperactive and Inattentive (Scales From K-SADS-E. Ordinal Raw Data Threshold Mx Modeling Approach Used)

Phenotype	Model	χ^2 (*df*)	$\lvert\Delta\chi^2\rvert$(*df* diff)	*p*	AIC
DSM-III-R	Saturated	1224.03(680)			
DSM-III-R	ACE	1224.52 (681)	.49(1)	.48	
DSM-III-R	AE	1224.58(682)	.06(1)	.81	–1.93
DSM-III-R	CE	1226.44(682)	1.92(1)	.17	–.08
DSM-III-R	E	1247.46(683)	22.88(2)	1.08e–05	18.6
K-SADS-E Hyperactive	Saturated	1592.9(680)			
K-SADS-E Hyperactive	ACE	1596.6 (684)	3.7(4)	.45	
K-SADS-E Hyperactive	AE	1597.6 (685)	1.0(1)	.32	–1.79
K-SADS-E Hyperactive	CE	1596.8 (685)	.2(1)	.65	–.95
K-SADS-E Hyperactive	E	1616.4(686)	19.8(2)	5.0e–05	15.9
K-SADS-E Inattentive	Saturated	1538.06(676)			
K-SADS-E Inattentive	ACE	1547.98 (682)	9.92(6)	.13	
K-SADS-E Inattentive	AE	1548.19 (683)	.21(1)	.65	–1.83
K-SADS-E Inattentive	CE	1549.6 (683)	1.62(1)	.20	–.178
K-SADS-E Inattentive	E	1582.2 (684)	34.22(2)	3.7e–08	17.4

Note: $N = 169$ for DZ twins and $N = 173$ for MZ.

(to DSM-III-R) prevalence of cases; we thus arrived at a prevalence of 7.1% for inattention and 6.1% for hyperactivity for retrospectively assessed childhood rates. Table 1 summarizes the Pearson correlation coefficients (calculated on total symptom counts) and the tetrachoric correlations and concordance rates for retrospectively assessed childhood ADHD.

Table 2 presents the results of fitting univariate biometric models to retrospectively assessed childhood symptom count data for each of the three phenotypes (DSM-III-R, K-SADS-E hyperactive and inattentive subscales) for both full and reduced models. Due to the extreme nonnormality of the data, a feature of Mx was employed which permits the analysis of categorical data, thus avoiding the limitations of nonnormality. The standardized parameter estimates and heritability for each variable (DSM-III-R, K-SADS-E hyperactive and K-SADS-E inattentive symptom counts) for the ACE symptom-based models are presented as well. The difference between the χ^2 of a general model (χ^2g) and that of a submodel (χ^2s) is itself a χ^2 with *df*(s) – *df*(g) degrees of freedom, thus the relative fit of submodels compared to the full ACE model can be tested.

As shown in Table 2, the model fitting for each of the three symptom count scales, DSM-III-R, K-SADS-E hyperactive, and K-SADS-E inattentive are quite similar. In each case, a full model that allowed for additive genetic, shared, and unique environmental influences to the phenotypic variances provided a satisfactory fit. The ACE models for DSM-III-R symptom count data accounted for the following components of variance: genetic = .29, 95% confidence intervals (CI) .00–.55; shared environmental = .11, 95% CI .00–.44; unique environmental = .60, 95% CI .45–.78; and K-SADS-E inattentive (genetic = .33, 95% CI .00–.61; shared environmental = .12, 95% CI .00–.48; unique environmental = .55, 95% CI .39–.73) and K-SADS-E Hyperactive (genetic = .30, 95% CI .00–.55; shared environmental = .09, 95% CI .00–.44; unique environmental = .61, 95% CI .48–.80) using retrospectively assessed childhood hyperactive and inattentive symptom count subscales.

For all three phenotypes, dropping either A or C failed to produce a model that was significantly worse (i.e., did not cause a significant degradation in fit), therefore, a model that does not allow for additive genetic contributions (CE) or shared environment (AE) is superior to the ACE model on the basis of parsimony. In each case, the values of the AIC criterion (Akaike, 1987) suggest that an AE model is superior, which is consistent with other researchers (Levy et al., 1997; Martin et al., 2002; Sherman, Iacono, et al., 1997). In all cases, removing both genetic (A) and shared environmental (C) terms from the model causes a statistically significant degradation of fit. These results demonstrate a clearly significant familial resemblance, however the relative importance of genes and family environment cannot be demonstrated unequivocally.

Not all the individual items of the K-SADS-E scales were found to be endorsed equally. Table 3 indicates the percentage of subjects among those who were given the repeat (adult perspective) interview and were considered affected using the 'broad diagnostic criteria' (49 subjects), who endorsed a particular item as applicable to them as children versus the percentage affected in the Mannuzza et al. (2002) sample.

In general, the endorsement patterns are fairly consistent, with the majority of the items on which there is overall agreement between the two samples (as indicated

https://doi.org/10.1375/twin.9.2.220 Published online by Cambridge University Press

Table 3

Per Cent of Subjects (n = 49) Considered 'Affected' Using Levy's Criteria Who Endorsed Each Symptom vs. % 'Affected' in Mannuzza et al. (2002) Study

Child Symptom	% endorsement Mannuzza sample	% symptom endorsement our sample	Child symptom variable Mannuzza sample	% endorsement our sample	% symptom endorsement
Difficulty w/attention	75	83.7*	Can't play quietly	22	30.6*
Doesn't listen	76	81.6*	Talks excessively	42	46.9*
Can't follow directions	34	83.7	Blurts out answers	46	28.6
Loses things	27	18.4*	Difficulty waiting turn	35	20.4
Easily distracted	84	75.5*	Interrupts or intrudes	24	38.8
Fidgets & squirms	84	75.5*	Acts before thinking	80	36.7
Can't remain seated	64	44.9	Shifts activities a lot	36	63.3

Note: Rows marked with * indicate agreement within 20% of the endorsement level found by Mannuzza et al., (2002).

by a * in Table 3) pertaining to the 'inattention' aspects of the disorder. This observation is consistent with work showing that inattentive symptoms tend to predominate in adult ADHD (Millstein et al., 1997; Zucker et al., 2002) and thus are perhaps more salient for the adult subject. We further found that by improvising a new scale comprised only of the frequently endorsed items (difficulty with attention, doesn't listen, can't follow directions, fidgets and squirms, can't remain seated, easily distracted), we achieved a heritability for this ad hoc symptom scale of .41 (95% CI .03–.54; with no indication of a common environmental effect) which, while still lower than usually reported, is statistically significant in the context of an ACE model and more consistent with some other published studies (Sherman, Iacono, et al., 1997).

At a phenotypic level, we calculated that the intraperson correlation between the retrospectively assessed phenotypes of K-SADS-E inattention and hyperactivity to be .66, comparable to that reported by Sherman, McGue, et al. (1997), who used third party informants. For these phenotypes we performed a bivariate genetic analysis to estimate the genetic and environmental contributions to the correlations. Results of model fitting are shown in Table 4, with the best fitting model (AE) highlighted. For the AE bivariate model, we calculated the variance due to genetic and unique environmental influences and the genetic and unique environmental correlations for hyperactivity and inattention symptom counts. The AE bivariate model (Figure 2) includes substantial genetic (r_G = .71; 95% CI .51–.88) and unique environmental (r_E =.63; 95% CI .46–.75) correlations between dimensions. Using this model, genes accounted

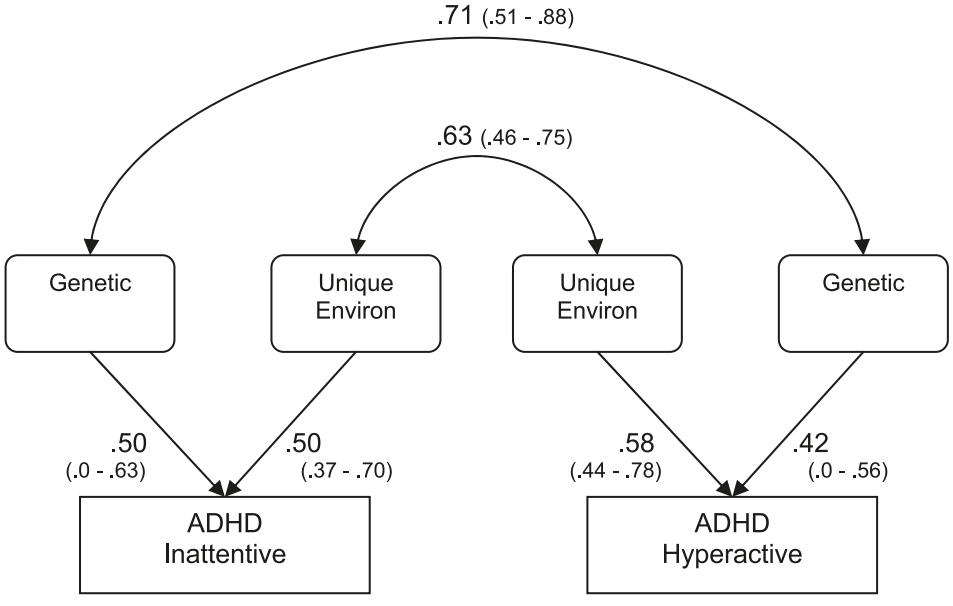

Figure 1

AE correlated factors view of the relationship between the inattentive and hyperactive symptom counts of ADHD.

Note: The numbers on the arrows between the latent factors (boxes with rounded corners) represent percentage of variance explained. Double-headed arrows represent the genetic and unique environmental correlations between latent factors.

https://doi.org/10.1375/twin.9.2.220 Published online by Cambridge University Press

Table 4

Model Fitting Results for Bivariate Cholesky Decomposition Using Total Number of Hyperactive Symptoms for ADHD (K-SADS-E) and Total Number of Inattentive Symptoms of ADHD (Also K-SADS-E)

Model tested (Hyperactivity and Inattentive Symptoms)	vs.	−2LL	*df*	$\lvert\Delta\chi^2\rvert$	Δ*df*	*p*	AIC
I. Saturated model		3107.4	1336				
II. ACE Cholesky Decomposition	I	3131.9	1356	24.6	20	< .10	
III. Dropping all covariance terms	II	3315.4	1359	183.4	3	< .01	177.4
IV. AE Cholesky Decomposition	**II**	**3132.3**	**1359**	**0.32**	**3**	**> .5**	**−5.67**
V. CE Cholesky Decomposition	II	3143.7	1359	11.8	3	> .05	
VI. E Cholesky	II	3174.3	1362	40.3	6	< .01	30.3

for 50% (95% CI .0–.63) and 42% (95% CI .0–.56) of the variance in inattentive and hyperactive symptom counts respectively, similar to the univariate estimates reported in Table 2, while unique environmental influences accounted for 50% (95% CI .37–.70) and 58% (95% CI .44–.78) of the variance in inattentive and hyperactive symptom counts, respectively.

Longitudinal Measures

As described above, subjects who indicated even mild ADHD symptoms in their retrospective assessment of themselves as children were asked to complete the same questions, evaluating their behavior as adults. The results of the paired sample *t* tests (retrospective child vs. current adult) on the inattentive and hyperactive symptom scores and the DSM-III-R symptom counts of those subjects who, as children, scored five or higher on the DSM-III-R criteria are presented in Table 4. Of the 49 subjects represented in the table, only 14 individuals were from twin pairs who scored concordantly above the 5+ inclusion threshold, thus we felt justified in ignoring the slight reduction in error variance caused by the few pairs (seven) in which both members of a twin pair were being included in the analysis.

When one member of those seven twin pairs was randomly excluded, resulting in 42 independent subjects, the results were identical in direction, quite similar in magnitude, and of equal significance. In addition, the results from the subset of subjects (20) who met the strict DSM-III-R criteria for childhood ADHD as children (8+ symptoms) were also similar in both direction and magnitude. Thus the subset of subjects who perceived themselves as having difficulty with ADHD-type behavior as children saw these behaviors significantly less troublesome as adults.

In addition to the drop in mean levels, Table 6 shows that the consistency of symptom persistence (childhood to adult) was fairly substantial and statistically significant. This indicates that subjects kept their self-perceived rank ordering over time (i.e., those that perceived themselves to be quite affected as children also thought so as adults while those that endorsed few symptoms as children also did so as adults); those who thought they were severely affected as children continued to perceive themselves as affected as adults, but less severely. These correlations (Table 6) were computed using only the 'A' twins to avoid the complications involved using correlated data.

Discussion

Our data suggest that while retrospective assessment of ADHD with an adult sample has much in common with results from children or adolescents, there are some differences. The first issue to be addressed is whether the prevalence of the disorder in our sample is comparable to other population-based studies. The prevalence rates found in our sample, for both childhood and adult ADHD, are lower than typically reported (Faraone et al., 2003), which could be due to several causes. For example, it is widely held (Kremen et al., unpublished manuscript) that population samples of veterans, in addition to being relatively homogeneous, are actually healthier than the general population. The filtering process of induction into the military removes many individuals (e.g., the grossly overweight, those with high blood pressure, and those

Table 5

Average Childhood Retrospective Ratings of ADHD Symptomatology vs. Current Adult Ratings for Probands Based on 5+ Symptom Criteria

Variable	Child $\bar{X}$	Adult $\bar{X}$	*N*	*SD*	*t*	Sig (2 tailed)
Inattention sx	13.31	7.57	49	3.02/2.89	11.07	< .001
Hyperactive sx	13.98	10.10	49	3.90/3.00	6.26	< .001
DSM-III-R rating	7.23	2.02	48	2.19/2.42	13.11	< .001

https://doi.org/10.1375/twin.9.2.220 Published online by Cambridge University Press

Table 6

Correlations Between Retrospectively Assessed Symptom Counts (Childhood) and Current Symptom Counts (Adult)

	N	Correlation	Sig.
DSM-III-R symptom counts Child vs. Adult	93	.252	.015
Childhood ADHD inattention sx & Adult ADHD inattention sx (K-SADS)	94	.259	.012
Childhood ADHD hyperactivity/impulsivity sx & Adult ADHD hyperactivity/impulsivity sx (K-SADS)	94	.296	.004

with incipient or early expressed mental disorders, or extremely low IQ) from the pool of inductees. These effects may still be noticeable many years later when military service is an ascertainment criterion, as is the case with the Vietnam Era Twin Registry sample. Another possible explanation is the existence of cohort differences in our sample; ADHD was rarely diagnosed during the 1960s, when these veterans would have been at the age of maximal risk.

Another possibility for the low prevalence rate as well as the relatively low observed heritability is the effect of retrospective recall. Perhaps the passage of time alone tends to diminish the ability of subjects to recall their childhood symptoms, resulting in an underestimation of disease prevalence and a larger measurement error effect. There is conflicting evidence for this hypothesis. While not assessed retrospectively, Danckaerts et al. (1999) found that using self-reports, adolescents with ADHD tended to underestimate their specific ADHD symptoms. Henry et al. (1994) reported that self-recollections correlated poorly with the same reports collected in childhood and with those of parents and teachers on the child (r = .04–.12). Some authors (Mannuzza et al., 2002; Shaffer, 1994) asserted that adults can not be considered valid informants of their own past behavioral patterns due to inaccuracies and difficulties in remembering early childhood signs and symptoms. Mannuzza suggested that based upon the test sensitivity of his results, in a population-based sample, only 27% of adults known to have ADHD when retrospectively diagnosed would be correctly identified. Barkley et al. (2002) suggested that self-recollections seem to underrepresent the severity of the childhood stage of the disorder, at least according to parents. Our subjects did not recall hyperactive and inattentive symptoms equally. Inattentive symptoms were recalled substantially more often ($p < .001$) which is consistent with Zucker et al. (2002) who found that adult probands tended to focus on the inattentive aspects of the disorder.

Not all studies are consistent with the hypothesis of forgetting deviant behavior with increasing age, however. Murphy and Schachar (2000) found that adults can give a reasonably accurate account of their childhood and current behavior, although interestingly, their subjects seemed to retrospectively inflate the severity of childhood symptoms compared to observers (their parents). If this finding were observed in our data, accurate childhood rates of ADHD might even be lower than observed. Existing data are fairly inconsistent on this point, suggesting that a prevalence measure alone is a poor indicator of methodological reliability.

A second dimension for estimating the validity of retrospective assessment is whether or not the retrospective data accurately reproduces the subtype structure of ADHD. For example, using the technique of latent class analysis, Todd and colleagues (2001) have identified several distinct subclasses within the domains of inattention, hyperactivity-impulsivity and combined-type ADHD. Our data are consistent with studies such as Todd's which suggest subclasses of ADHD within the primary diagnosis. We found a phenotypic correlation of .66 between the K-SADS-E defined subclasses of hyperactive/impulsive and inattentive ADHD, consistent with the results of Sherman, McGue, et al. (1997). At a phenotypic level, our results are consistent with prevailing results (Bauermeister et al., 1992) and theory (American Psychiatric Association, 1994). In addition, our bivariate modeling of childhood symptom counts (K-SADS-E hyperactivity and inattention) demonstrated a reasonably high genetic correlation (r = .60), suggesting that the dimensions of hyperactivity and inattentiveness, while distinct, are significantly associated. This result is consistent with the Todd et al. (2001) characterization of the ADHD phenotype as consisting of severe hyperactive, severe inattentive and combined classes.

Most twin studies indicate that ADHD symptoms are highly heritable, with as much as 70% to 80% of phenotypic variation being accounted for by genetic factors (e.g., Eaves et al., 1997; Faraone, 2004a; Martin et al., 2002; Thapar et al., 2000), although this finding has not been universal. Martin et al. (2002), for example, found that self-rated scores from adolescents on the Strengths and Difficulties Questionnaire (Goodman, 1997) resulted in equal correlations in MZ and DZ twins which caused the most acceptable biometric model to be one that included zero heritability. This result was based on only 286 pairs of adolescent twins, however. The heritabilities that we found using ordinal symptom count data were in the .3 to .4 range, which, while substantially below most estimates, were still statistically significant.

Importantly, we also found that the individual scale items endorsed retrospectively most frequently, were very consistent with other studies (Mannuzza et

https://doi.org/10.1375/twin.9.2.220 Published online by Cambridge University Press

al., 2002; Ward et al., 1992); another indicator of the comparability of our approach with previous work. Mannuzza et al. (2002) found six symptoms with high sensitivity and specificity (distractibility, concentration difficulties, complaints of inattention, acting before thinking, being 'on the go', and fidgeting and squirming) that compare favorably to the symptoms we found to be most frequently endorsed as well as being most heritable. These same symptoms were found by Ward et al. (1992) to best discriminate adults previously diagnosed with ADHD from normal controls. We further found that by creating a new scale comprised only of the frequently endorsed items (difficulty with attention, doesn't listen, can't follow directions, fidgets and squirms, can't remain seated, easily distracted), we achieved a more comparable heritability, consistent with other published literature using third-party informants. Sherman, McGue, et al. (1997), for example, found a heritability of only 39% for teacher ratings using the Teacher's Rating Form. Since the E variance component includes error of measurement, the inflation of E would be expected if subjects, due to retrospective assessment, had more difficulty recalling their symptoms reliably; this is consistent with our finding that by eliminating the infrequently endorsed items, our observed heritability goes up, suggestive of more reliable measurement.

Zucker et al. (2002) suggested an explanation for the inconsistencies and low heritabilities produced by retrospective reporting. They suggested that memories necessary to rate ADHD symptoms experienced in childhood are organized as semantic rather than episodic memories: 'The task of responding to a questionnaire about ADHD symptoms experienced in childhood does not draw on one's memory of a single episode or salient life event' (p. 387). This type of memory is represented as general memories for a prototypical experience of repeated and/or frequent symptoms (Kessler et al., 1999). As such, the constructed nature of such semantic memories may also impact their validity.

Some researchers argue that retrospective diagnoses of childhood ADHD made on the basis of self-reports, particularly for an older adult, would, in most cases, be difficult to establish (e.g., Jackson & Farrugia, 1997; Shaffer, 1994). They argue that inaccuracies and difficulties in remembering early childhood signs and symptoms are probable (Mannuzza et al., 2002). We do not dispute the plausibility of this position, but the pattern of our results in several key dimensions are consistent with other studies (Martin et al., 2002; Sherman, McGue, et al., 1997; Ward, 1992; Zucker et al., 2002) which show a modest to high heritability and a substantial decline in symptoms as the proband ages.

Consequently, we conclude that while assessing ADHD in adult probands may be less accurate than with children or adolescents, since it shares several characteristics with other assessment techniques (lower heritability, but nevertheless, demonstrable familiality, similar item endorsement patterns, strong association between hyperactive and inattentive subtypes, and a longitudinal decline in symptom severity) it remains a viable diagnostic strategy, even with population samples. Heritability estimates ultimately derive from differences in the covariance between MZ and DZ twins, a statistic known to be attenuated by decreases in reliability; it seems reasonable that symptom reports obtained 30 to 40 years after the time period in question may result in lower estimates of heritability. Prevalence estimates lower than many others that have been reported again may reflect memory distortions due to the long time interval since the relevant symptoms occurred; it may also reflect a sample that was screened for physical and mental health at the time of military induction and may have excluded a number of individuals with more severe ADHD symptoms or associated problems (e.g., problems with the law).

Potential limitations of our data include retrospective reporting of childhood symptoms by our middle-aged participants over a lengthy time period and an exclusively male sample that may have been healthier than the population. Consequently, the results may not generalize to females or nonveteran males.

References

Akaike, H. (1987). Factor analysis and AIC. *Psychometrika, 52,* 317–332.

American Psychiatric Association. (1987). *Diagnostic and statistical manual of mental disorders* (3rd ed., rev.). Washington, DC: Author.

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

Barkley, R. (1990). *Primary symptoms and conceptualization in attention deficit hyperactivity disorder: A handbook for diagnosis and treatment*. New York: Guilford Press.

Barkley, R. (1998). *Attention deficit hyperactivity disorder: A handbook for diagnosis and treatment*. New York: Guilford Press.

Barkley, R. A., Fischer M., Smallish L., & Fletcher, K. (2002). The persistence of attention-deficit/hyperactivity disorder into young adulthood as a function of reporting source and definition of disorder. *Journal of Abnormal Psychology, 111*, 279–289.

Bauermeister, J. J., Algeria, M., Bird, H. R., Rubio-Stipec, M., & Canino, G. (1992). Are attentional-hyperactivity deficits unidimensional or multidimensional syndromes? Empirical findings from a community survey. *Journal of the American Academy of Child and Adolescent Psychiatry*, *31*, 423–431.

Biederman, J., Faraone, S. V., Milberger, S., Curtis, S., Chen, L., Marrs, A., Ouellette, C., Moore, P., & Spencer, T. (1996). Predictors of persistence and remission of ADHD: Results from a four year prospective

https://doi.org/10.1375/twin.9.2.220 Published online by Cambridge University Press

follow-up study of ADHD children. *Journal of the American Academy Child and Adolescent Psychiatry, 35*, 343–351.

Biederman, J., Mick, E., & Faraone, S. V. (2000). Age-dependent decline of symptoms of attention deficit hyperactivity disorder: Impact of remission definition and symptom type. *American Journal of Psychiatry, 157*, 816–818.

Chen, P. Y., & Popovich, P. M. (1995). *Correlation: Parametric and non-parametric measures*. Thousand Oaks, CA: SAGE Publications.

Danckaerts, M., Heptinstall, E., Chadwick, O., & Taylor, E. (1999). Self-report of attention deficit and hyperactivity disorder in adolescents. *Psychopathology, 32*, 81–92.

Derks, E. M., Dolan, C. V., & Boomsma, D. I. (2004). Effects of censoring on parameter estimates and power in genetic modeling. *Twin Research*, 7, 659–669.

Dobson, A. (2002). *An introduction to generalized linear models*. Boca Raton, FL: Chapman & Hall/CRC.

Eaves, L. J., Silberg, J. L., Meyer, J. M., Maes, H. H., Simonoff, E., Pickles, A., Rutter, M., Neale, M. C., Reynolds, C. A., Erikson, M. T., Heath, A. C., Loeber, R., Truett, K. R., & Hewitt, J. K. (1997). Genetics and developmental psychopathology: 2. The main effects of genes and environment on behavioral problems in the Virginia Twin Study of Adolescent Behavioral Development. *Journal of Child Psychology and Psychiatry, 38*, 965–980.

Eisen, S., Neuman, R., Goldberg, J., & Rice, T. W. (1989). Determining zygosity in the Vietnam Era Twin Registry: An approach using questionnaires. *Clinical Genetics, 35*, 423–432.

Eisen, S., True, W., Goldberg, J., Henderson, W., & Robinette, C. D. (1987). The Vietnam Era Twin (VET) Registry: Method of construction. *Acta Geneticae Medicae et Gemellologiae, 36*, 61–66.

Emory University. (1993). *Diagnostic Interview Schedule for Children — Parent version* (PC-DISC). Atlanta: Author.

Falconer, D. S. (1965). The inheritance of liability to certain diseases estimated from the incidence among relatives. *Annals of Human Genetics, 29*, 51–76.

Faraone, S. V. (2004a). Genetics of adult attention deficit disorder. *The Psychiatric Clinics of North America*, *27*, 303–321.

Faraone, S. V. (2004b). Etiology and pathophysiology of adult attention deficit hyperactivity disorder. *Primary Psychiatry, 11*, 28–40.

Faraone, S. V., & Biederman, J. (1994). Is attention deficit hyperactivity disorder familial? *Harvard Review Psychiatry, 1*, 271–287.

Faraone, S. V., Biederman, J., & Friedman, D. B. A. (2000). Validity of DSM-IV subtypes of attention-deficit/hyperactivity disorder: A family study perspective. *Child and Adolescent Psychiatry, 39*, 300–307.

Faraone, S., Biederman, J., & Mick, E. (in press). The age dependent decline of attention-deficit/hyperactivity disorder: A meta-analysis of follow-up studies. *Psychological Medicine*.

Faraone, S. V., Perlis, R. H., Doyle A. E., Smoller, J. W., Goralnick, J. J., Holmgren, M. A., & Sklar, P. (2005). Molecular genetics of attention-deficit/hyperactivity disorder. *Biological Psychiatry*, *57*, 1313–1323.

Faraone, S. V., Sergeant, J., Gillberg, C., & Biederman, J. (2003). The worldwide prevalence of ADHD: Is it an American condition? *World Psychiatry, 2*, 104–113.

Gillis, J. J., Gilger, J. W., Pennington, B. F., & DeFries, J. C. (1992). Attention deficit disorder in reading-disabled twins: Evidence for a genetic etiology. *Journal of Abnormal and Child Psychology, 20*, 303–315.

Goodman, R. (1997). The Strengths and Difficulties Questionnaire: A research note. *Journal of Child Psychology and Psychiatry*, *38*, 581–586.

Goodman, R., & Stevenson, J. (1989). A twin study of hyperactivity. I. An examination of hyperactivity scores and categories derived from Rutter teacher and parent questionnaires. *Journal of Child Psychology and Psychiatry, 30*, 671–689.

Heiligenstein, E., Conyers, L. M., Berns, A. R., Miller, M. A., & Smith, M. A. (1998). Preliminary normative data on DSM-IV ADHD in college students. *Journal of the American College of Health*, *46*, 185–188.

Henry, B., Moffitt, T. E., Caspi A., Langley, J., & Silva, P. A. (1994). On the 'remembrance of things past': A longitudinal evaluation of the retrospective method. *Psychological Assessment, 6*, 92–101.

Jackson, B., & Farrugia, D. (1997). Diagnosis and treatment of adults with attention deficit hyperactivity disorder. *Journal of Counseling and Development*, *75*, 55–62.

Kessler, R. C., Adler, L., Barkley, R., Biederman, J., Conners, C. K., Demler, O., Faraone, S. V., Greenhill, L. L., Howes, M. J., Secnik, K., Spencer, T., Ustun, T. B., Walters, E. E., & Zaslavsky A. M. (in press). The prevalence and correlates of adult ADHD in the United States: Results from the national comorbidity survey replication. *The American Journal of Psychiatry*.

Kessler, R. C., Mroczek, D. K., & Belli, R. F. (1999). Retrospective adult assessment of childhood psychopathology. In D. Schaffer, C. P. Lucas, & J. E. Richters (Eds.), *Diagnostic assessment in child and adolescent psychopathology* (pp. 256–284). New York: Guilford Press.

Kremen, W. S., Lyons, M. J., Boake, C., Xian, H., Jacobson, C., Waterman, B., Eisen, S. A., Goldberg, J., Faraone, S. V., Tsuang, M. T. *A discordant twin study of premorbid cognitive ability in schizophrenia*. Unpublished manuscript, Department of Psychiatry, University of California at San Diego, La Jolla, CA.

Levy, F., Hay, D. A., McStephen, M., Wood, C., & Waldman, I. (1997). Attention deficit hyperactivity disorder: A category or a continuum? Genetic analysis

https://doi.org/10.1375/twin.9.2.220 Published online by Cambridge University Press

of a large-scale twin study. *Journal of the American Academy of Child and Adolescent Psychiatry, 36,* 737–744.

Mannuzza, S., Klein, R. G., Bessler, A., Malloy, P., & LaPadula, M. (1993). Adult outcome of hyperactive boys: Educational achievement, occupational rank, and psychiatric status. *Archives of General Psychiatry, 50*, 565–576.

Mannuzza, S., Klein, R. G., Klein, D. F., Bessler, A., & Shrout, P. (2002). Accuracy of adult recall of childhood attention deficit hyperactivity disorder. *The American Journal of Psychiatry, 159*, 1882–1888.

Martin, N., Scourfield, J., & McGuffin, P. (2002). Observer effects and heritability of childhood attention-deficit hyperactivity disorder symptoms. *British Journal of Psychiatry, 180*, 260–265.

McGue, M. (1992). When assessing twin concordance, use the probandwise not the pairwise rate. *Schizophrenic Bulletin*, *18*, 171–176.

Millstein, R. B., Wilens, T. E., Biederman, J., & Spencer, T. J. (1997). Presenting ADHD symptoms and subtypes in clinically referred adult with ADHD. *Journal of Attention Disorders*, *2,* 159–166.

Murphy, K., & Barkley R. A. (1996). Attention deficit hyperactivity disorder adults: Comorbidities and adaptive impairments. *Comprehensive Psychiatry*, *37,* 393–401.

Murphy, P., & Schachar, R. (2000). Use of self-ratings in the assessment of symptoms of attention deficit hyperactivity disorder in adults. *American Journal Psychiatry, 157*, 1156–1159.

Nadder, T. S., Silberg, J. L., Rutter, M., Maes, H. H., & Eaves, L. J. (2001). Comparison of multiple measures of ADHD symptoms: A multivariate genetic analysis. *Journal of Child Psychology and Psychiatry, 42*, 475–486.

Neale, M. C., Boker, S. M., Xie, G., & Maes, H. H. (2002). *Mx: Statistical modeling* (6th ed.). Richmond, VA: Department of Psychiatry, Medical College of Virginia.

Neale, M. C., Eaves, L. J., & Kendler, K. S. (1994). The power of the classical twin study to resolve variation in threshold traits. *Behavior Genetics, 24,* 239–258.

Neuman, R. J., Todd, R. D., Heath, A. C., Reich, W., Hudziak, J. J., Bucholz, K. K., Madden, P. A., Begleiter, H., Porjesz, B., Kuperman, S., Hesselbrock, V., & Reich, T. (1999). Evaluation of ADHD typology in three contrasting samples: A latent class approach. *Journal of the American Academy of Child and Adolescent Psychiatry, 38*, 25–33.

Orvaschel, H., & Puig-Antich, J. (1987). *Schedule for affective disorder and schizophrenia for school-age children: Epidemiologic 4th Version.* Fort Lauderdale, FL: Nova University for Psychological Study.

Rasmussen, E. R., Neuman, R. J., Heath, A. C., Levy, F., Hay, D. A., & Todd, R. D. (2002). Replication of the latent class structure of attention-deficit/hyperactivity disorder (ADHD) subtypes in a sample of Australian twins. *Journal of Child Psychiatry and Psychology, 43*, 1018–1028.

Shaffer, D. (1994). Attention deficit hyperactivity disorder in adults. *The American Journal of Psychiatry, 151*, 633–638.

Sharp, W. S., Gottesman, R. F., Greenstein, D. K., Ebens, C. L., Rapoport, J. L., & Castellanos, F. X. (2003). Monozygotic twins discordant for attention-deficit/hyperactivity disorder: Ascertainment and clinical characteristics. *Journal of the American Academy of Child and Adolescent Psychiatry, 42*, 93–97.

Sherman, D. K., Iacono, W. G., & McGue, M. K. (1997). Attention-deficit hyperactivity disorder dimensions: A twin study of inattention and impulsivity-hyperactivity. *Journal of the American Academy of Child and Adolescent Psychiatry, 36,* 745–753.

Sherman, D. K., McGue, M. K., & Iacono, W. G. (1997). Twin concordance for attention deficit hyperactivity disorder: A comparison of teachers' and mothers' reports. *American Journal of Psychiatry, 154*, 532–535.

Seidman, L. J., Beiderman, J., Weber, W., Hatch, M., & Faraone, S. V. (1998). Neuropsychological function in adults with attention-deficit hyperactivity disorder. *Biological Psychiatry, 44*, 260–268.

Thapar, A., Harrington, R., Ross, K., & McGuffin, P. (2000). Does the definition of ADHD affect heritability? *Journal of the American Academy of Child and Adolescent Psychiatry, 39*, 1528–1536.

Thapar, A., Holmes, J., Poulton, K., & Harrington, R. (1999). Genetic basis of attention deficit and hyperactivity. *British Journal of Psychiatry*, *174,* 105–111.

Thissen, D., Pommerich, M., Billeaud, K., & Williams, V. S. L. (1995). Item response theory for scores on tests including polytomous items with ordered responses. *Applied Psychological Measurement, 19,* 39–49.

Todd, R. D. (2000). Genetics of attention deficit/hyperactivity disorder: Are we ready for molecular genetic studies? *American Journal of Medical Genetics, 96*, 241–243.

Todd, R. D., Rasmussen, E. R., Neuman, R. J., Reich, W., Hudziak, J. J., Bucholz, K. E., Madden P. A. F., & Heath, A. (2001). Familiality and heritability of subtypes of attention deficit hyperactivity disorder in a population sample of adolescent female twins. *American Journal of Psychiatry, 158*, 1891–1898.

Tsuang, M. T., Bar, J. L., Harley, R. M., & Lyons, M. J. (2001). The Harvard twin study of substance abuse: What we have learned. *Harvard Review of Psychiatry, 9,* 267–279.

van den Oord, E. J. C. C., Simonoff, E., Eaves, L. J., Pickles, A., Silberg, J., & Maes, H. (2000). An evaluation of different approaches for behavior genetic analyses with psychiatric symptom scores. *Behavior Genetics, 30,* 1–18.

https://doi.org/10.1375/twin.9.2.220 Published online by Cambridge University Press

van den Oord, E. J. C. C., & van der Ark, L. A. (1997). A note on the use of the Tobit approach for tests scores with floor or ceiling effects. *British Journal of Mathematical and Statistical Psychology, 50*, 351–364.

Ward, M. F., Wender, P. H., & Reimherr F. W. (1992). The Wender Utah rating scale: An aid in theretrospective diagnosis of childhood attention deficit hyperactivity disorder. *American Journal of Psychiatry, 150*, 885–888.

Zucker, M., Morris, M. K., Ingram, S. M., Morris, R. D., & Bakeman, R. (2002). Concordance of self- and informant ratings of adults' current and childhood attention-deficit/hyperactivity disorder symptoms. *Psychological Assessment, 14*, 379–389.

https://doi.org/10.1375/twin.9.2.220 Published online by Cambridge University Press