Strong genetic effects on cross-situational antisocial behaviour among 5-year-old children according to mothers, teachers, examiner-observers, and twins' self-reports

Louise Arseneault,[1] Terrie E. Moffitt,[1,2] Avshalom Caspi,[1,2] Alan Taylor,[1] Fruhling V. Rijsdijk,[1] Sara R. Jaffee,[1,2] Jennifer C. Ablow,[3] and Jeffrey R. Measelle[3]

[1]Social, Genetic, and Developmental Psychiatry Research Centre, Institute of Psychiatry, King's College London, UK; [2]Department of Psychology, University of Wisconsin, Madison, USA; [3]Department of Psychology, University of Oregon, Eugene, USA

Background: Early childhood antisocial behaviour is a strong prognostic indicator for poor adult mental health. Thus, information about its etiology is needed. Genetic etiology is unknown because most research with young children focuses on environmental risk factors, and the few existing studies of young twins used only mothers' reports of behaviour, which may be biased. **Method:** We investigated genetic influences on antisocial behaviour in a representative-plus-high-risk sample of 1116 pairs of 5-year-old twins using data from four independent sources: mothers, teachers, examiner-observers previously unacquainted with the children, and the children themselves. **Results:** Children's antisocial behaviour was reliably measured by all four informants; no bias was detected in mothers', teachers', examiners', or children's reports. Variation in antisocial behaviour that was agreed upon by all informants, and thus was pervasive across settings, was influenced by genetic factors (82%) and experiences specific to each child (18%). Variation in antisocial behaviour that was specific to each informant was meaningful variation, as it was also influenced by genetic factors (from 33% for the children's report to 71% for the teachers' report). **Conclusions:** This study and four others of very young twins show that genetic risks contribute strongly to population variation in antisocial behaviour that emerges in early childhood. In contrast, genetic risk is known to be relatively modest for adolescent antisocial behaviour, suggesting that the early-childhood form has a distinct etiology, particularly if it is pervasive across situations. **Keywords:** Antisocial behaviour, environmental influences, genetics, self-reports, twins.

Children who exhibit antisocial behaviour at ages 3 to 5 years are significantly more likely to display conduct problems and delinquency in adolescence, to meet diagnostic criteria for antisocial personality disorder in adulthood, and to be recidivistic and violent criminals (Tremblay, Pihl, Vitaro, & Dobkin, 1994; Caspi, Moffitt, Newman, & Silva, 1996; Stevenson & Goodman, 2001). Although some children exhibiting antisocial behaviour do not grow up to have antisocial personalities, virtually all antisocial children suffer adult maladjustment (Robins, 1966; Farrington, Gallagher, Morley, St. Ledger, & West, 1988; Moffitt, Caspi, Harrington, & Milne, 2002). Given the strong prognostic significance of early childhood antisocial behaviour for life-course development, it is critical to understand its etiology (Keenan & Wakschlag, 2002).

Most etiological studies of young children's antisocial behaviour have focused on environmental risk factors, such as poverty or ineffective parenting (Shaw, Owens, Vondra, Keenan, & Winslow, 1996; Haapasalo & Tremblay, 1994; Keenan, 2001). The role of genetic influences has seldom been considered. Until recently, quantitative genetic studies of antisocial behaviour have focused on adolescent and adult samples to the virtual exclusion of pre-school children (Rhee & Waldman, 2002). This is surprising given the prognostic significance of childhood-onset antisocial behaviour.

Two opposing predictions are offered about the genetic and environmental origins of childhood-onset antisocial behaviour. One developmental theory suggests that environmental factors are strongest in early childhood, but then give way to increasing genetic influences as children grow older and are able to select for themselves environments that are correlated with their genotypes (Scarr & McCartney, 1983). This theory would predict that genetic influences might increase after childhood when adolescents who are genetically predisposed to antisocial activities shape their environment by choosing deviant peers who will encourage their antisocial behaviour (Rowe & Osgood, 1984). This theory that genetic effects are weakest in childhood is known to apply to cognitive abilities. Developmental behavioural genetics research has revealed that environmental influences account for most of the variation observed in cognitive abilities during early childhood, but genetic influence actually increases as children develop from childhood to

adulthood (McGue, Bouchard, Iacono, & Lykken, 1993; Plomin, Fulker, Corley, & DeFries, 1997).

In contrast to this theory predicting weak genetic influences for early relative to later antisocial behaviour, a taxonomic theory of antisocial development has suggested the opposite; that antisocial behaviour evident in early childhood should have strong genetic influences (Moffitt, 1993; DiLalla & Gottesman, 1989). According to this theory, antisocial behaviour that begins in childhood has its origins in neurodevelopmental vulnerabilities known to be heritable (e.g., undercontrolled temperament, language delay), and is likely to become life-course persistent. In contrast, antisocial behaviour that begins in adolescence is relatively transient and has its origins in peer social processes. This developmental theory would thus predict strong heritability for early childhood samples. The goal of the present study is to test these two competing hypotheses by examining antisocial behaviour among 5-year-old children.

Increasing evidence suggests that genetic factors do play an important role in children's early-emerging antisocial behaviour (Schmitz, Fulker, & Mrazek, 1995; van den Oord, Verhulst, & Boomsma, 1996; van den Oord, Boomsma, & Verhulst, 2000; van der Valk, Verhulst, Stroet, & Boomsma, 1998; Dionne, Tremblay, Boivin, Laplante, & Pérusse, 2003; van der Valk, van den Oord, Verhulst, & Boomsma, 2001). However, to date, no epidemiological study has been able to establish the genetic and environmental origins of antisocial behaviour in early childhood without being constrained by the use of parent reports only. Alternative sources of information about children's behaviour problems are hard to come by: pre-school children are generally thought to be unable to report on their own behaviour problems (Edelbrock, Costello, Dulcan, Kalas, & Conover, 1985; Boyle et al., 1993; Schwab-Stone, Fallon, Briggs, & Crowther, 1994); reports from teachers are usually not available because, in most countries, children under age 5 years have not started school; and observational measures are difficult and costly to collect in the context of the large samples required by behavioural genetic research designs. As a result, in genetic epidemiology, mothers (and in two studies, fathers contributed to a parental rating score) are the sole informants about pre-school children's antisocial behaviour. This reliance on mothers' reports may be problematic because concerns have been raised about the reliability and validity of mothers' reports. First, mothers' reports may be biased by mothers' own psychopathology or criminal history (Dumas & Wekerle, 1995; Fergusson, Lynskey, & Horwood, 1993; Chilcoat & Bresleau, 1997). Second, mothers may have limited information about what is normative for young children. Third, mothers' reporting skill is subject to variation in their intellectual abilities and educational background, which are known to be correlated

with their children's behavioural problems (Patterson & Yoerger, 1997). These concerns have led some child development experts to claim that mothers' reports of behaviour problems are of limited value because these parental measures are suffused by multiple forms of rating bias (Kagan, 1998). Although this claim may be a slight exaggeration of the known facts (Simonoff et al., 1995, 1998; van der Valk et al., 2001), it is certainly true that knowledge about the genetic and environmental origins of childhood-onset antisocial behaviour is currently restricted to knowledge about the genetics of parental ratings.

In the present epidemiological twin study, we investigate genetic and environmental influences on 5-year-old children's antisocial behaviour, as assessed by four different sources of information: the children's mothers, their teachers, examiner-observers previously unacquainted with the children, and the children themselves.

Method

Sample

Participants are members of the Environmental Risk (E-risk) Longitudinal Twin Study, which investigates genetic and environmental factors shaping children's development. The E-risk sampling frame was two consecutive birth cohorts (1994 and 1995) in the Twins' Early Development Study (TEDS), a birth register of twins born in England and Wales (Trouton, Spinath, & Plomin, 2002). The full register is administered by the government's Office of National Statistics (ONS), which invited parents of all twins born in 1994–95 to enrol. Of 15,906 twin pairs born in these two years, 71% joined the register. Our sampling frame excluded opposite-sex twin pairs and began with the 73% of register families having same-sex twins.

The E-risk Study sought a sample size of 1100 families to allow for attrition in future years of the longitudinal study while retaining statistical power. An initial list of 1210 families was drawn from the register to target for home visits, a 10% oversample to allow for nonparticipation. The probability sample was drawn using a high-risk stratification strategy. High-risk families were those in which the mother had her first birth when she was 20 years of age or younger. We used this sampling to replace high-risk families who were selectively lost to the register via non-response and to ensure sufficient base rates of problem behaviours given the low base rates expected for 5-year-old children. Early first childbearing was used as the risk-stratification variable because it was present for virtually all families in the register, it is relatively free of measurement error, and it is a known risk factor for children's problem behaviours (Maynard, 1997; Moffitt & the E-Risk Study Team, 2002). In the final sample, two-thirds of Study mothers accurately represent all mothers in the general population (aged 15–48) in England and Wales in 1994–95. The other one-third of Study mothers (younger only) constitute a 160% oversample of mothers who were at high risk based on their

young age at first birth (15–20 years). To ensure that the findings in this article represent unbiased estimates of the general population, the data were weighted for all analyses. This weighting makes the proportion of young mothers in the sample equivalent to the overall proportion in the population (Birth Statistics, 1996).

Of the 1210 families targeted, 7 were discovered to be ineligible for inclusion in our study because the twins had moved overseas, did not speak English, were being reared by neither biological parent, or were opposite-sex. Of the 1203 eligible families, 1116 (93%) participated in home-visit assessments when the twins were age 5 years. The sample includes 56% monozygotic (MZ) and 44% dizygotic (DZ) twin pairs. Sex is evenly distributed within zygosity (49% male). Data were collected within 120 days of the twins' fifth birthday. Research workers visited each home for 2.5 to 3 hours, in teams of two. While one interviewed the mother, the other tested the twins in sequence in a different part of the house. Families were given shopping vouchers for their participation, and children were given colouring books and stickers. All research workers had university degrees in behavioural science, and experience in psychology, anthropology, or nursing. With parents' permission, questionnaires were posted to the children's teachers, and teachers returned questionnaires for 94% of cohort children.

A substantial proportion of children in the E-risk sample showed disruptive behaviour at age 5; consistent with DSM-IV (American Psychiatric Association, 1994) diagnostic criteria, children with 3 or more symptoms could be given a diagnosis of conduct disorder (unweighted, the prevalence of conduct disorder in the sample was 8.5%; weighted to represent the population, it was 6.6%). Within this group, a smaller number of children with 5 or more symptoms met criteria for severe conduct disorder (unweighted prevalence = 3.4%, weighted to represent the population, it was 2.5%).

Antisocial behaviour

The present study successfully avoided problems caused by shared method variance across sources (Bank, Dishion, Skinner, & Patterson, 1990): one examiner interviewed the mother, another examiner interviewed and rated the children, the teacher returned a postal questionnaire, and a coder who did not test the children scored their videotaped self-reports.

Mothers' reports. Mothers' reports of children's antisocial behaviour were obtained in interviews using the Child Behavior Checklist (CBCL; Achenbach, 1991a). For this study focusing on antisocial behaviour, we used the Delinquency and Aggression scales supplemented with DSM-IV (American Psychiatric Association, 1994) items assessing conduct and oppositional defiant disorder (e.g., 'spiteful, tries to get revenge', 'uses force to take something from another child'). Scores ranged from 0 to 72 ($M = 15.56$, $SD = 11.40$). The internal consistency reliability of the mother report was .92.

Teachers' reports. Teachers' reports of antisocial behaviour were obtained using the Teacher Report Form (TRF; Achenbach, 1991b) supplemented as above. The behaviour of both twins was rated by the same teacher for 79% of the twins. Scores ranged from 0 to 74 ($M = 5.66$, $SD = 9.10$). The internal consistency reliability of the teacher report was .95.

Examiners' observations. After the home visit, examiners rated each twin on the Dunedin Behavioural Observation Scale, which includes 9 items measuring disruptive behaviour (e.g., hostility, lability, roughness) (Caspi, Henry, McGee, Moffitt, & Silva, 1995). Each behaviour was defined in explicit terms, and the examiner evaluated whether each characteristic was observed (0) not at all, (1) somewhat, or (2) definitely. A version if this instrument was initially used in the American Collaborative Study on Cerebral Palsy, Mental Retardation, and Other Neurological Disorders of Infancy and Childhood (Goldsmith & Gottesman, 1981) and later modified for use with 3- and 5-year-old children in the Dunedin Multidisciplinary Health and Development Study. Follow-up of the Dunedin sample showed that these disruptive behaviours predicted adolescent behaviour problems, as well as adult psychopathology (Caspi, 2000). Scores ranged from 0 to 18 ($M = 2.22$, $SD = 3.46$). The internal consistency reliability of the examiner report was .90 and the inter-rater reliability coefficient was .70.

Children's self-reports. We used the Berkeley Puppet Interview (BPI) to obtain self-reports from the twins about their own antisocial behaviour (Measelle, Ablow, Cowan, & Cowan, 1998). The BPI is a developmentally-appropriate instrument designed to investigate symptomatology in 4- to 8-year-old children. The BPI was administered to each twin separately. In the BPI, the examiner introduces two identical fluffy animal puppets (Iggy and Ziggy) to the child, and the puppets invite the child to join them in a conversation in which they tell the child things about themselves and the child tells them about him/herself. The two puppets make opposite statements (e.g., Iggy: 'I hit kids a lot' – Ziggy: 'I don't hit kids') in a counterbalanced order. The child is then asked to tell the puppets how he/she behaves. Children are allowed to respond verbally or non-verbally by pointing or touching the puppet to indicate their answer. All examiners completed a one-week certification-training course designed by Ablow and Measelle (Ablow & Measelle, 1999).

Children were administered 19 items covering three BPI scales that assess antisocial behaviour: Overt Aggression/Hostility (e.g., 'I fight with other kids'), Conduct Problems (e.g., 'I take things that don't belong to me'), and Oppositionality (e.g., 'I don't do what my teacher asks me to do'). All interviews were videotaped to score the children's answers later. Each item was coded on a 7-point Likert scale ranging from 1 (no symptom) to 7 (definite symptom). Two different coders scored each interview, with inter-rater reliability exceeding .90 for all coders. Scores ranged from 31 to 106 ($M = 51.48$, $SD = 13.38$), and the internal consistency reliability was .82. Further psychometric evaluation conducted as part of the McArthur Research Network on Psychopathology and Development attests to the reliability and validity of the BPI in this age group (Ablow et al., 1999). Data for the BPI were missing for 353 children leaving valid data for both twins in 901 pairs,

81% of sample families. Missingness was caused by the child not being able to complete the interview, by the examiner's and/or coder's assessment that the child did not understand the task, or by disruption or lack of privacy for the interview.

Statistical methods

Genetic model fitting. We used maximum likelihood estimation of model parameters in univariate and multivariate genetic models of children's antisocial behaviour (Neale & Cardon, 1992; Plomin, DeFries, McClearn, & McGuffin, 2001). When one phenotypic measure is analysed, models decompose variance in children's antisocial behaviour into latent additive genetic (A; i.e., the sum of the average effects of individual alleles at all loci), latent shared environmental (C), and latent nonshared environmental (E) factors. In cases where DZ correlations are less than half the MZ correlations, it is possible to examine two alternative models. First, a genetic dominance (D; i.e., interaction effects between alleles on the same locus) factor can be estimated instead of a C factor (ADE model). Second, a sibling interaction model (AE_S) can be fitted to the data to estimate genetic and environmental parameters free of sibling interaction effects (Thapar, Holmes, Poulton, & Harrington, 1999). When we analysed all four measures of antisocial behaviour simultaneously, we tested three models: a *biometric model*, a *psychometric model*, and a *rater-bias model* (Hewitt, Silberg, Neale, Eaves, & Erickson, 1992; van der Valk et al., 2001; van den Oord et al., 2000). Testing and comparing these three models enables us to examine the extent of the consensus among the four informants, and whether information unique to each informant represents measurement error, potential bias, or valuable data in the four measures.

The *biometric model* (Figure 1) partitions the variances and the covariances of the four measures of children's antisocial behaviour into two components: (1) common variance, which represents the variance that is shared by all four measures (conceptualised as the consensus between informants); and (2) unique variance, which represents the variance that is not shared between the four measures (conceptualised as information that is unique to each informant). This model posits that variation common to the four measures of antisocial behaviour is influenced by a set of 'common' genetic and environmental factors and, in addition, that the variation unique to each four measures is also influenced by 'unique' genetic and environmental factors. This model assumes that each informant reports information that is shared or common with other informants, but also reports unique, valuable information for which genetic and environmental factors can be estimated.

In contrast to the biometric model, the *psychometric model* (Figure 2) posits that genetic and environmental factors influencing variation common to the four measures do not influence the measures directly; rather, they influence each of the four measures of antisocial behaviour via a latent factor (Figure 2). This factor represents the underlying common variation between informants that indexes individual differences in the liability to antisocial behaviour that is pervasive across

settings. In other words, this model isolates from the measured variance of each report, the consensus across informants, and thus represents antisocial behaviour agreed upon by the informants. This model is more parsimonious than the biometric model because it estimates fewer parameters.

The *rater-bias model* (Figure 3) partitions the variances and the covariances of the four measures of children's antisocial behaviour into three components: (1) reliable trait variance which represents the variance, free of rater bias, that is shared by the four informants (conceptualised as the reliable consensus across informants); (2) rater bias, which represents the correlated errors across twins, within each informant; and (3) unreliability due to measurement error, which represents the residuals that are not correlated across twins. In contrast to both the biometric and psychometric models, the rater-bias model assumes that only the variance that is shared across informants is valuable and informative. It posits that this reliable, agreed-upon trait variance is influenced by a set of common genetic and environmental factors. The rest of the variance is accounted for by unreliability that is partitioned into a rater-bias component and a measurement-error component. The rater-bias model is more parsimonious than either the biometric or psychometric models because it estimates fewer parameters.

Goodness-of-fit indices. The goal of fitting different structural equations to twin data is to account for the observed covariance structure using fewest possible parameters. In comparing the fit of different models, we used four model-selection statistics to have a general picture of the model fit because overall fit statistics are affected by large sample size (Bollen, 1989). The first was the χ^2 goodness-of-fit statistic. Large values compared to model degrees of freedom indicate poor model fit to the observed covariance structure. When two models are nested (i.e., identical with the exception of constraints placed on the sub-model), the difference in fit between them can be evaluated with the χ^2 difference, using as its degrees of freedom the *df* difference from the two models. When the χ^2 difference is not statistically significant, the more parsimonious model is selected, as the test indicates that additional constraints do not improve the model fit. The second model-selection statistic was Akaike's Information Criterion (AIC) (Burnham & Anderson, 1998). The AIC is founded on ideas from information theory and provides a goodness-of-fit measure that penalises models for increasing complexity and can be used with non-nested models. When comparing two models, the model with the lowest AIC value is selected as the best fitting model. The third model-selection statistic was the Root Mean Square Error of Approximation (RMSEA) which is an index of the model discrepancy, per degree of freedom, from the observed covariance structure (MacCallum, Browne, & Sugawara, 1996). A RMSEA of less than or equal to .06 indicates a good fitting model (Hu & Bentler, 1999). The fourth model-selection statistic was the Bayesian Information Criterion (BIC), where increasingly negative values correspond to increasingly better-fitting models. In comparing two models, differences of BIC between 6 and 10 give strong evidence in favour of the model with the smaller value (Raftery,
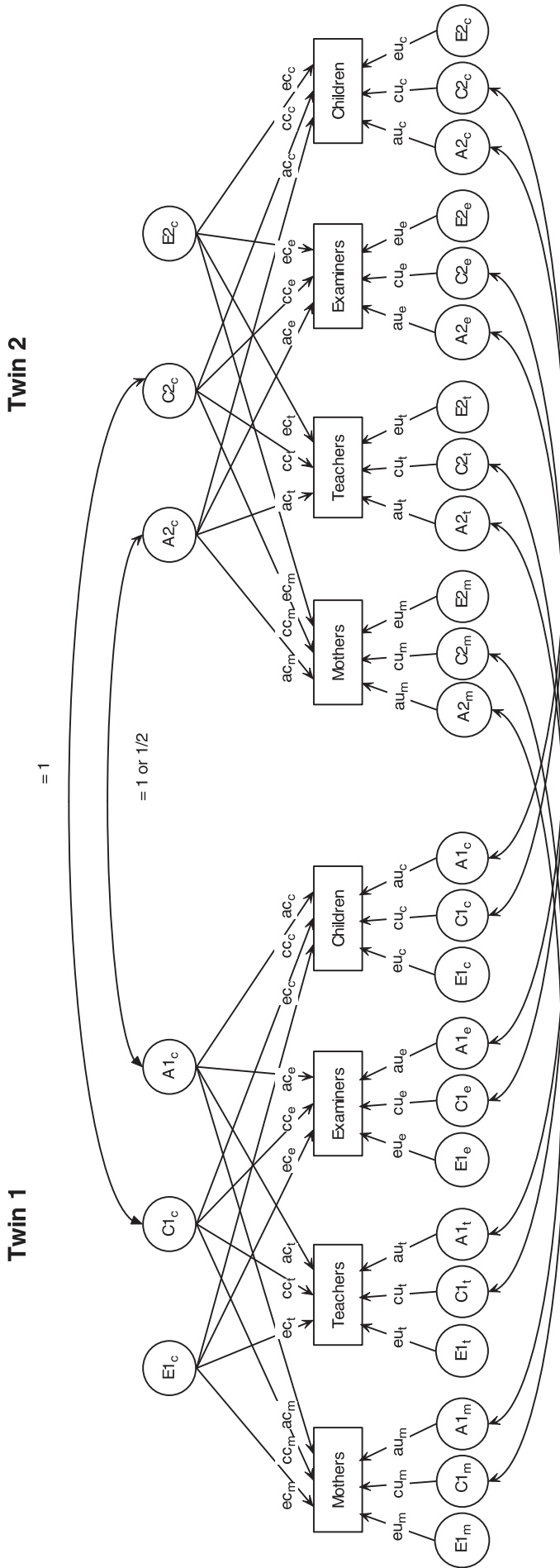
Figure 1 The biometric model for antisocial behaviour according to mothers', teachers', examiner-observers', and children's self-reports. Measured variables are depicted by rectangles and latent variables by circles. The model is separated into two parts; the upper part (above the measured variables) contains the effects on the variance common to all four informants – common genetic (A1$_c$, where 1 stands for twin1), common shared environment (C1$_c$), and common non-shared environment (E1$_c$) – and the lower part (below the measured variables) contains the effects on the variance unique to each informant; for example with the mothers' ratings – unique genetic (A1$_m$ where m stands for mothers' report), unique shared environment (C1$_m$), and unique non-shared environment (E1$_m$). Single-headed paths are the effects of latent variables on measured variables (e.g., the path ac$_m$ represents the effect of the common genetic factor on mothers' ratings of children's antisocial behaviour). Double-headed paths represent correlations between latent variables which indicate, by standard biometric genetic theory, that MZ twins totally share both common and unique genetic influences ($r = 1.0$), that DZ twins share only half of both common and unique genetic influences ($r = .5$), and that both MZ and DZ twins completely share common and unique shared-environmental influences. Variances of all latent variables were fixed to one for model identification
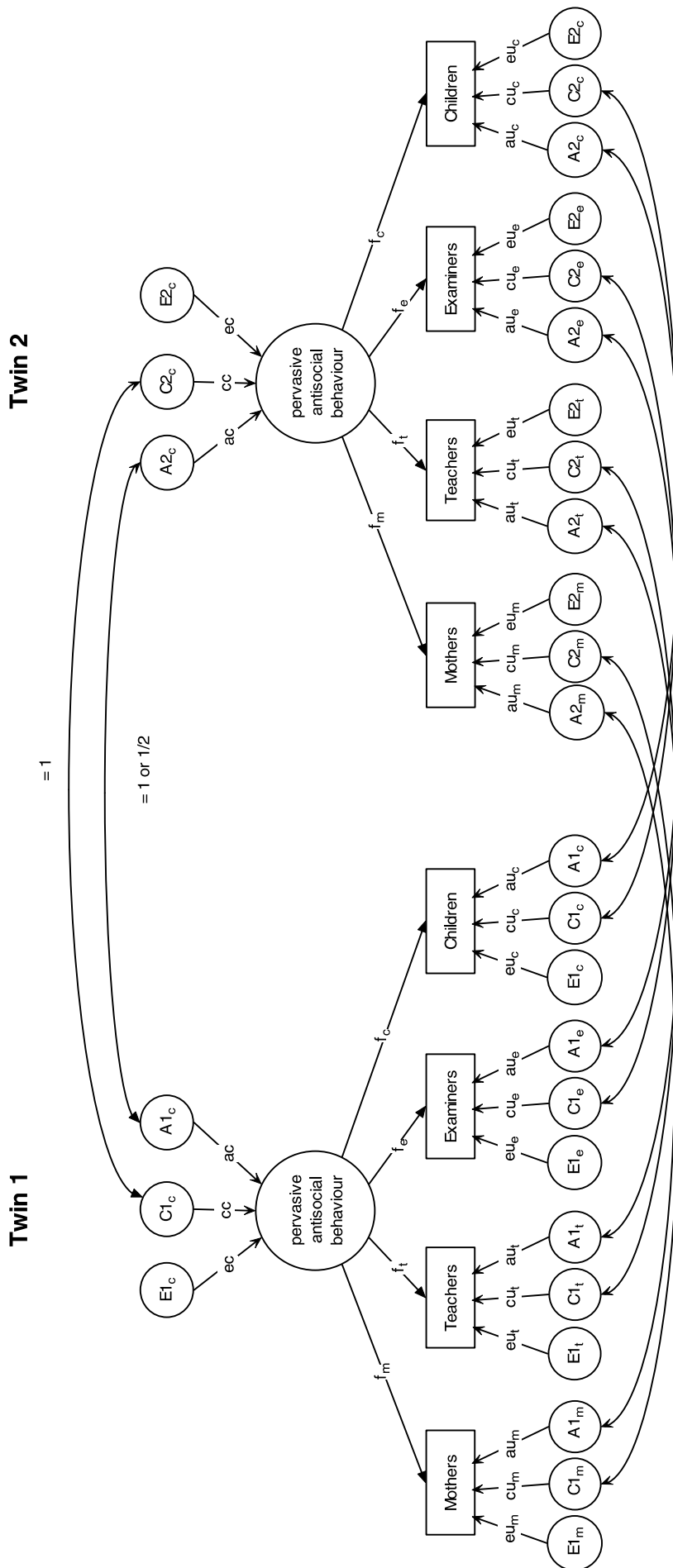
Figure 2 The psychometric model for antisocial behaviour according to mothers', teachers', examiner-observers', and children's self-reports. For the psychometric model, we used the same conventions as for the biometric model with the only difference being the presence of a common factor of children's antisocial behaviour capturing the shared variance between the four informants. A1$_c$, C1$_c$, and E1$_c$ (where *1* stands for twin 1) are genetic and environmental estimates for this common factor that represents a liability to antisocial behaviour. The single-headed paths – f_m, f_t, f_e, f_c – are the factor loadings of each of the informants on the common factor. The lower part of this model is identical to the biometric model
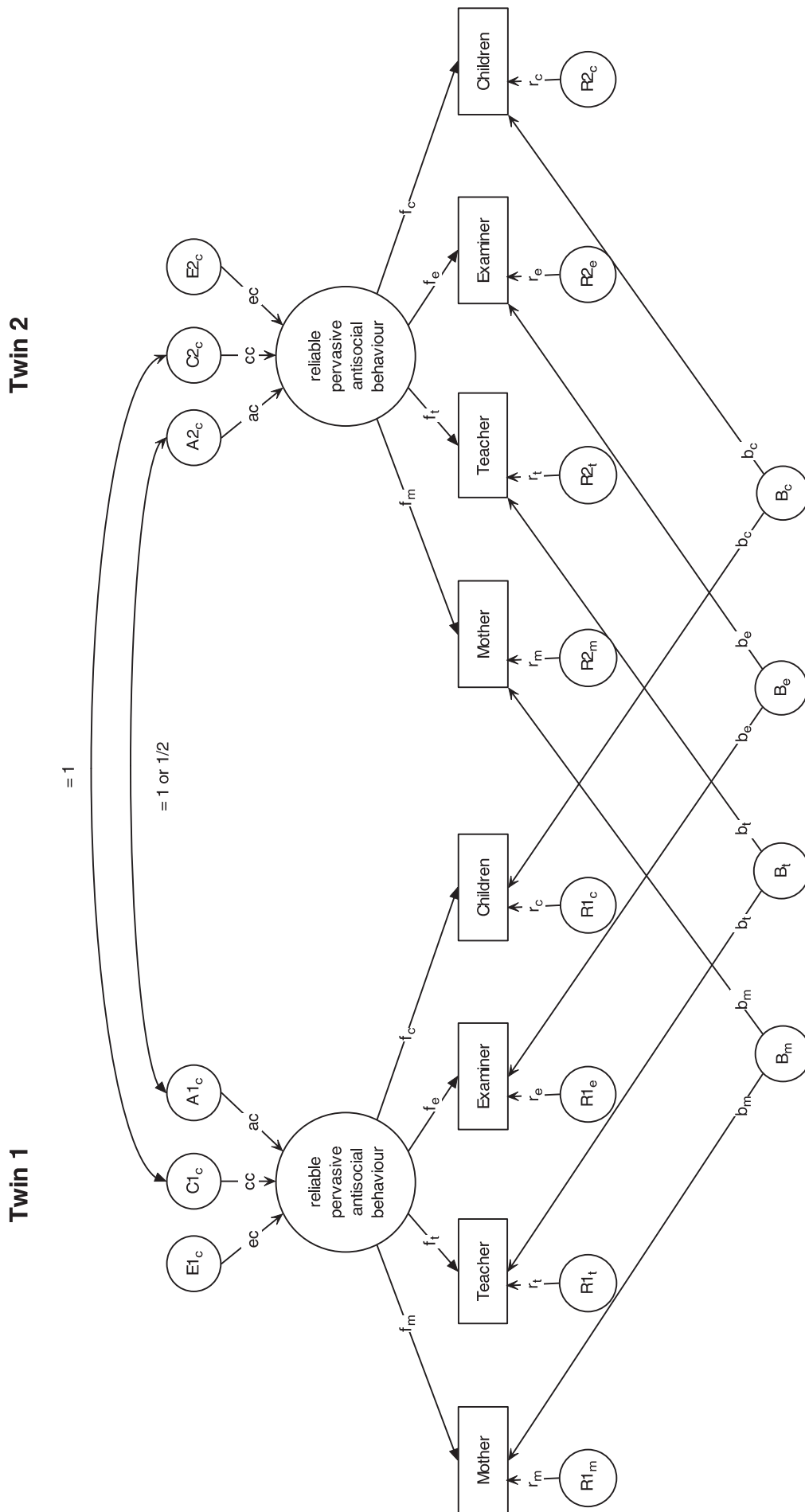
Figure 3 The rater bias model for antisocial behaviour according to mothers', teachers', examiner-observers', and children's self-reports. For the rater-bias model, we used the same convention as for the psychometric model with the difference being in the lower part of the model. The set of unique estimates is composed of the rater bias estimates (e.g., B_m where m stands for mothers' report) and their corresponding residuals (e.g., $R1_m$). The single-headed paths (e.g., r_m and b_m) are the effects of the rater bias and the residuals estimates on measured variables. The upper part of this model is identical to the psychometric model

1995). We used BIC as our preferred model comparison statistic because it tends to favour more parsimonious models than AIC, especially for larger sample sizes, and also because it gives information on the degree to which a model fits better compared to other models. The psychometric and the rater-bias models cannot be compared to the biometric model by using the χ^2 difference because these two models are not nested within the biometric model. In this case, the BIC was used to compare the fit of the different models. The rater-bias model can be compared to the psychometric model using the χ^2 difference because it is nested within it.

To reduce the skewness of antisocial behaviour distributions, we applied a square-root transformation to all four antisocial behaviour scales. To correct for sex differences in mean levels of antisocial behaviour, we standardised all four measures within sex. To provide unbiased estimates that can be generalised to the population, we weighted all statistics so the sample represents the proportion of young mothers in the UK (Birth Statistics, 1996). Structural equation modelling programs Mplus (Muthén & Muthén, 1998) and Mx (Neale, 1997) were used to analyse data for this study.

Results

Descriptive results

Twin resemblance was seen across all four measures of antisocial behaviour (Table 1). Correlations between MZ twins were substantial, ranging from .43 (for children's self-reports of antisocial behaviour) to .79 (for teachers' reports of antisocial behaviour). Correlations between DZ twins were modest, ranging from .24 (for children's self-reports) to .35 (for mothers' reports). These correlations were not affected by the square-root transformation of the scales (Table 1). The significant correlations for both MZ and DZ twins established that there is familial aggregation for antisocial behaviour. The differences between the MZ and DZ correlations provided rough estimates of the heritability of antisocial behaviour, derived by calculating 2(r MZ—r DZ) where r is the within-pair correlation. As shown in Table 1, the MZ correlation for each measure of antisocial behaviour was approximately twice the size of the DZ correlation, suggesting substantial additive genetic effects on individual differences in children's antisocial behaviour. The importance of child-specific environmental effects is reflected in the difference between the MZ twin correlation and 1.0, because the MZ twin correlation theoretically represents all twin resemblance attributable to genetic and shared environmental influences.

We observed significant sex differences in mean levels of antisocial behaviour (Table 1). Across all four measures, male children exhibited more antisocial behaviour, regardless of zygosity. The effect sizes (d) associated with these sex differences were small (ranging from .30 to .36), which is consistent with previous research (Moffitt, Caspi, Rutter, & Silva, 2001). Despite sex differences in mean levels of

Table 1 Correlations between MZ and DZ twins for mothers', teachers', examiner-observers', and children's self-reports of antisocial behaviour

Informants N pairs/Maximum pairs	Mothers' report 1109/1116		Teachers' report 1035/1116		Examiners' report 1103/1116		Children's report 901/1116	
	MZ	DZ	MZ	DZ	MZ	DZ	MZ	DZ
r for untransformed antisocial behaviour	.67	.35	.79	.27	.63	.32	.43	.24
r (95% CI) for square root antisocial behaviour	.70 (.66–.74)	.38 (.30–.46)	.78 (.75–.81)	.31 (.23–.39)	.61 (.56–.66)	.35 (.28–.43)	.42 (.35–.49)	.24 (.15–.33)
M*	15.31	15.84	6.06	5.22	2.27	2.17	51.55	51.40
(SD)	(11.25)	(11.58)	(10.11)	(7.81)	(3.50)	(3.42)	(13.67)	(13.05)
N pairs	611	498	565	470	609	494	490	411
	M / F	M / F	M / F	M / F	M / F	M / F	M / F	M / F
r (95% CI) for square root antisocial behaviour — M	.70 (.64–.76)	.41 (.31–.52)	.83 (.79–.86)	.27 (.15–.39)	.65 (.58–.72)	.32 (.21–.44)	.41 (.30–.52)	.21 (.08–.35)
r (95% CI) for square root antisocial behaviour — F	.70 (.64–.75)	.34 (.24–.45)	.74 (.69–.79)	.27 (.15–.39)	.57 (.50–.65)	.39 (.29–.50)	.43 (.33–.53)	.27 (.15–.39)
M — M	16.66	17.74	7.65	6.67	2.89	2.83	53.75	53.07
M — F	14.02	13.93	4.51	3.78	1.68	1.50	49.45	49.86
(SD) — M	(12.12)	(12.77)	(11.76)	(9.16)	(3.84)	(3.90)	(14.54)	(13.56)
(SD) — F	(10.20)	(9.88)	(7.89)	(5.88)	(3.03)	(2.72)	(12.42)	(12.38)
N pairs — M	296	248	278	230	295	243	234	193
N pairs — F	315	250	287	240	314	251	256	218

MZ = Monozygotic twin pairs; DZ = Dizygotic twin pairs; M = Male; F = Female; CI = Confidence Intervals.
*Means (M) and standard deviation (SD) are reported for the untransformed antisocial behaviour variables.

antisocial behaviour, there were no significant sex differences in the MZ or DZ twin correlations (as shown by the overlapping confidence intervals of the correlations). This suggests that genetic and environmental influences on antisocial behaviour were of the same magnitude for the two sexes.

Univariate model fitting

Looking at each informant's report separately, model-fitting revealed three noteworthy results (Table 2). First, it revealed the important contribution of genetic factors to antisocial behaviour among 5-year-old children. Second, it revealed the significant contribution of child-specific environmental experiences to children's antisocial behaviour. Third, it revealed only a small contribution of shared-environmental experiences to children's antisocial behaviour accounting for a maximum of 12% of the variation, which would require a much larger sample for statistical significance. For all four measures of antisocial behaviour, more parsimonious AE models provided a better fit to the data than full ACE models. Because DZ twin correlations were less than half the MZ twin correlations for the teachers' reports, we also fitted a model testing a dominance factor (ADE model) and a sibling interaction-effects model (AE$_s$ model), but these models did not fit any better than the AE model. Depending on the measure examined, genetic influences accounted for between 42% (in children's self-reports) and 69% (in mothers' reports)

of the variance in children's antisocial behaviour. Child-specific environmental effects (plus measurement error) accounted for between 24% and 39% of the variance in mother, teacher, and examiner reports of antisocial behaviour. Child-specific environmental effects accounted for 58% of the variance in children's self-reports of their antisocial behaviour, possibly reflecting the greater measurement error inherent in 5-year-old children's self-reports or children's awareness of their own behaviour across settings.

Correlations among the four informants and multivariate model fitting

The correlations between the four measures of antisocial behaviour ranged from .14 to .28 (Table 3). These correlations may seem low but are in keeping with cross-informant correlations reported in previous studies of children's psychiatric problems (Achenbach, McConaughy, & Howell, 1987; van der Ende, 1999); meta-analyses reveal that the mean parent-teacher correlation is .28 (we report .28) and the mean correlation between subjects themselves and adult informants is .22 (we report between .18 and .21). A principal axis factor analysis yielded one factor, with eigenvalues greater than 1.0, suggesting that each of the four measures is a valid, but imperfect indicator of a liability to antisocial behaviour.

Fit statistics appeared to be affected by our large sample size. However, the psychometric model had

Table 2 Univariate estimates of genetic and environmental contributions to antisocial behaviour for mothers', teachers', examiner-observers' and children's self-reports

| | Variance components | | | Overall model fit | | | | | | Model difference test | | |
| | Genetic | Environmental | | | | | | | | | | |
	h^2 (95% CI)	Shared c^2 (95%CI)	Non-shared e^2 (95% CI)	χ^2	df	p	AIC	RMSEA	BIC	$\Delta\chi^2$	Δdf	p
Mother-report (N = 1109)												
ACE	.61 (.46–.72)	.08 (.00–.22)	.31 (.27–.35)	.840	3	.840	−5.160	.000	−20.194			
AE	**.69 (.66–.73)**		**.31 (.27–.34)**	**1.922**	**4**	**.750**	**−6.078**	**.000**	**−26.123**	**1.081**	**1**	**.298**
Teacher-report (N = 1035)												
ACE	.76 (.64–.79)	.00 (.00–.11)	.24 (.22–.28)	16.674	3	.001	10.674	.094	−4.152			
AE	**.76 (.72–.79)**		**.24 (.22–.28)**	**16.674**	**4**	**.002**	**8.674**	**.078**	**−11.095**	**.000**	**1**	**1**
ADE[†]	.62 (.27–.78)	.13 (.00–.49)	.24(.21–.28)	16.076	3	.001	10.076	.092	−4.750			
AE$_s$[‡]	.75 (.69–.80)	.01 (−.04–.06)	.25(.20–.31)	16.599	3	.001	10.599	.078	−4.227			
Examiner-report (N = 1103)												
ACE	.48 (.32–.64)	.12 (.00–.27)	.40 (.35–.45)	.836	3	.841	−5.164	.000	−20.181			
AE	**.61 (.56–.65)**		**.39 (.35–.44)**	**2.970**	**4**	**.563**	**−5.030**	**.006**	**−25.053**	*2.134*	1	.144
Child-report (N = 901)												
ACE	.30 (.08–.47)	.10 (.00–.29)	.59 (.53–.67)	3.037	3	.386	−2.963	.009	−17.374			
AE	**.42 (.35–.48)**		**.58 (.52–.65)**	**4.079**	**4**	**.395**	**−3.921**	**.007**	**−23.135**	**1.042**	**1**	**.307**

N refers to the number of twin pairs with complete data.
AIC = Akaike's Information Criterion; RMSEA = Root Mean Square Error of Approximation; BIC = Bayesian Information Criterion.
CI = Confidence Intervals.
A = h^2 = genetic variance component; C = c^2 = shared environment variance component; E = e^2 = non-shared environment variance; D = d^2 = dominance variance component; s = sibling interaction effect
[†]D effect is in the second column; [‡]s effect is in the second column.

Table 3 Internal consistency reliability and correlations between mothers', teachers', examiner-observers', and children's self-reports of antisocial behaviour

	Informant			
	Mothers	Teachers	Examiners	Children
Mothers	.92			
Teachers	.28	.95		
Examiners	.14	.21	.90	
Children	.18	.21	.20	.82

Internal consistency reliability coefficients (Cronbach alpha) are reported in the diagonal.
N children = 1680.

the best overall combination of fit to the data and parsimony (Table 4). The biometric model had a smaller chi-square and AIC than the other models, and an equivalent RMSEA. However, it is not as parsimonious as the psychometric model which also showed a fair fit to the data as indicated by all model-selection statistics, and especially by the large negative BIC (the difference in BIC between the two models was greater than 10). Next, we tested whether the full psychometric model could be simplified by eliminating various parameters of the model. Table 4 shows that dropping the common C parameter and the four unique C parameters resulted in the most parsimonious model.

The results of the best-fitting model are shown in Table 5. The best-fitting psychometric model revealed that children's antisocial behaviour was reliably measured by four different informants, as indicated by factor loadings on each of the four measures ranging from .39 to .56 (Table 5, 1st row; 21% of the variance of the measures was explained by a common latent factor). The variation in the common latent factor of children's antisocial beha-

viour, representing the consensus among informants, was influenced by genetic factors (82%) and by child-specific environmental factors (18%). In addition, the variation in children's antisocial behaviour that was uniquely reported by each informant was not simply measurement error; rather, it represented meaningful variation between children and it reflected the influence of unique genetic factors (ranging from 28% of the total variance of children's self-reports to 51% of the total variance of mothers' reports; Table 5, 7th row) and unique child-specific environmental factors including measurement error (ranging from 56% for children's self-reports to 20% for teachers' reports; Table 5, 8th row). Parallels between common and unique genetic estimates were established by comparing genetic estimates of the common latent factor (82%) to the proportion of the unique variance accounted for by genetic factors, ranging from 33% (for the children's self-reports) to 71% (for the teacher's reports) (Table 5, 9th row). This comparison shows that genetic factors greatly influenced variation in antisocial behaviour that was common to the four informants, and also variation that was specific to each informant (to a lesser extent the children's reports). Further sensitivity analyses were conducted by re-estimating the psychometric model using raw data rather than standardised, transformed variables; the results were unchanged (these results are available upon request, as are all analysis scripts).

When looking back at each informant's report separately, we observed no significant sex differences in the twin correlations (Table 1). Nevertheless, we tested a sex-specific version of the best-fitting multivariate model (Neale et al., 1992). The sex-specific version fit better than a sex-independent model ($\chi^2 = 80.74$, $df = 59$, $p = .03$). Closer

Table 4 Fit indices of multivariate models for mothers', teachers', examiner-observers', and children's self-reports of antisocial behaviour

	Overall model fit						Model difference test		
	χ^2	df	p	AIC	RMSEA	BIC	$\Delta\chi^2$	Δdf	p
Overall model									
Biometric model	72.497	48	.013	−23.503	.035	−250.706			
Psychometric model	**92.489**	**54**	**.001**	**−15.511**	**.040**	**−271.115**			
Rater bias model	161.843	58	.000	43.843	.065	−228.694	69.354[a]	4	.000
Simplification of best fitting model[b]									
No common A	121.123	55	.000	11.123	.053	−249.214	28.634	1	.000
No common C	92.489	55	.001	−17.511	.040	−277.848	.000	1	1
No unique As[c]	161.843	58	.000	45.843	.065	−228.694	69.354	4	.000
No unique Cs	99.012	58	.001	−16.988	.040	−291.525	6.523	4	.163
No common A + no unique As	243.133	59	.000	125.133	.085	−154.138	150.644	5	.000
No common C + no unique Cs	**99.012**	**59**	**.001**	**−18.988**	**.039**	**−298.259**	**6.523**	**5**	**.259**

N = 840 twin pairs with complete data on all different data sources.
A = genetic variance component; C = shared environment variance component.
AIC = Akaike's Information Criterion; RMSEA = Root Mean Square Error of Approximation; BIC = Bayesian Information Criterion.
[a]The chi-square difference is between the psychometric and the rater bias models.
[b]The full psychometric model is the comparison model.
[c]This constrained model is the equivalent of the rater bias model.

Table 5 Common and unique parameter estimates of genetic and environmental contributions to antisocial behaviour accounting for the common underlying trait, the total measured variance, and the unique measured variance as determined by the psychometric model[a]

	Latent factor	Of the common underlying trait			
		Mothers	Teachers	Examiners	Children
1. Factor loading		.47 (.40–.54)	.56 (.48–.64)	.39 (.31–.46)	.40 (.32–.47)
2. Factor loading squared[b]		.22 (.16–.29)	.31 (.23–.41)	.15 (.10–.21)	.16 (.11–.22)
3. Common genetic	.82 (.74–.89)				
4. Common non-shared environment	.18 (.11–.26)				
		Of the total measured variance			
5. Common genetic		.18 (.13–.24)	.26 (.18–.34)	.12 (.08–.17)	.13 (.09–.18)
6. Common non-shared environment		.04 (.02–.06)	.06 (.03–.08)	.03 (.01–.04)	.03 (.02–.05)
7. Unique genetic		.51 (.44–.57)	.49 (.40–.57)	.47 (.41–.54)	.28 (.20–.36)
8. Unique non-shared environment		.27 (.23–.32)	.20 (.16–.24)	.38 (.33–.43)	.56 (.49–.63)
		Of the unique measured variance[c]			
9. Unique genetic		.65 (.60–.70)	.71 (.65–.76)	.56 (.49–.61)	.33 (.25–.41)
10. Unique non-shared environment		.35 (.30–.40)	.29 (.24–.35)	.44 (.39–.51)	.66 (.59–.75)

$N = 840$ twin pairs with complete data on all different data sources.

[a]We obtained similar estimates and model fit statistics when fitting the psychometric model using unstandardised/untransformed data

[b]Factor loadings squared represent the proportion of variance in each measure that is accounted for by the liability to antisocial behaviour (latent common factor).

[c]To get the proportion of the unique variance of antisocial behaviour that is accounted for by genetic and non-shared environmental factors, we calculated, for each informant, the proportion of the total variance that was unique to each informant by subtracting from each measure's total variance the variance accounted for by the latent factor. Then, we divided the genetic and environmental estimates we obtained from the psychometric model by the proportion of unique variance. Using the mothers' findings as an example: $1-.22 = .78$ (proportion of the total variance of the mothers' report that is unique to this informant; $.51/.78 = 65\%$ of the variance unique to the mothers' report was influenced by genetic factors.

inspection revealed that sex differences were limited to one data source, the examiners' observations: the unique genetic effect associated with examiner ratings of antisocial behaviour (parameter au_e in Figure 2) was significantly higher among boys ($au_e = .57$; 95% CI: .49–.66) than girls ($au_e = .38$; 95% CI: .28–.47) and the unique nonshared environmental effect of these ratings (eu_e) was higher among girls ($eu_e = .47$; 95% CI: .39–.56) than boys ($eu_e = .26$; 95% CI: .21–.33). The genetic and environmental influences on the common factor of antisocial behaviour (ac and ec), and the factor loadings (f_m, f_t, f_e, f_c), did not differ by sex.

Discussion

Antisocial behaviour that begins early in life is a strong risk factor for poor mental health and criminality later in life. Using mothers', teachers', examiner-observers', and children's self-reports of antisocial behaviour, this study established that this risk factor is highly heritable. Our finding of strong heritability for age-5 antisocial behaviour is not consistent with developmental theory posing that genetic influences are weakest in childhood, when family rearing environment influences behaviour and before individuals choose their own niches in life. Although we have observed relatively weak genetic and notable common environmental influences on age-5 IQ in the E-risk sample (Kuntsi &

Eley, unpublished work), the opposite pattern was found here for antisocial behaviour problems. The observed heritability estimate of 82% for early-onset antisocial behaviour agreed upon by multiple informants approaches the heritability estimates for the most heritable psychiatric disorders: schizophrenia, autism, and hyperactivity (McGuffin, Riley, & Plomin, 2001). Findings also suggested that the information provided by all four informants (mothers, teachers, examiner-observers, and 5-year-old children) is valuable for research uses because the four reports of children's antisocial behaviour included reliable and non-biased information that was agreed upon by all informants. Moreover, each source provided its own unique information that appeared to be meaningful information because it was accounted for by genetic factors, not mere measurement error.

Limitations to our study must be mentioned. First, our analyses relied on a basic assumption underlying all twin studies, the 'equal environments assumption' that MZ and DZ twins had the same exposure to environmental influences that make twins similar on antisocial behaviour (Kendler, 1993). However, MZ twins' families are no worse than DZ twins' families on parenting risks associated with antisocial behaviour, and ways in which MZ twins are treated more similarly are not associated with antisocial risk (DiLalla, 2002). Second, we did not model parental assortative mating for antisocial behaviour, although it is known to exist (Krueger,

Caspi, Moffitt, Bleske, & Silva, 1998). However, assortative mating acts to inflate shared-environment effects, which were very small in our study, suggesting that such inflation is not problematic. Third, we assume that findings can be generalised from twins to the population of singletons (Rutter, Thorpe, Greenwood, Northstone, & Golding, 2003). This assumption is probably defensible because twin–singleton comparisons find no notable differences in behaviour problems or personality (Johnson, Krueger, Bouchard, & McGue, 2002; Moilanen et al., 1999; Gjone & Novik, 1995; Kendler, Martin, Heath, & Eaves, 1995; Levy, Hay, McLaughlin, Wood, & Waldman, 1996; Simonoff et al., 1997; van den Oord, Koot, Boomsma, Verhulst, & Orleveke, 1995). Fourth, like all twin studies, we assumed that the effects of genetic and environmental factors were additive so that no gene–environment interactions or gene–environment correlations were operating (Rutter & Silberg, 2002). This is unlikely to be the case; the heritability coefficient may conceal instances in which children's genetic liability toward antisocial behaviour influences the kinds of environments they encounter and moderates the effects of those environments. In fact, our theory of how childhood-onset antisocial behaviour develops into life-course persistent antisocial behaviour emphasises these processes of gene–environment interplay (Moffitt, 1993) and the aim of future research in the E-risk Study is to test for correlations and interactions between genes and environments. Fifth, although we could verify that there was no difference in the magnitude of heritability of antisocial behaviour for boys and girls, we cannot test whether different genes were influencing their behaviour because funding limitations precluded sampling opposite sex-twins. Sixth, we selected the best-fitting model (the psychometric model) based on parsimony, but we could not eliminate completely the possibility that the biometric model could be correct as well. In either case, the conclusion would remain the same: that pervasive childhood antisocial behaviour is highly heritable. Seventh, our findings from multiple informants need replication using larger samples to detect potentially significant shared-environmental effects, but there are very few large twin studies with data from multiple informants. Limitations aside, the findings of this study have implications for theory about etiology, for methodology in future research, and for the work of clinicians.

The genetic contribution we observed for our single-informant measures (42% for children, 61% for examiner-observers, 69% for mothers, and 76% for teachers) is supported by similar high estimates from four other studies relying on single measures of antisocial behaviour in large representative samples of very young twins. Dionne et al. (2003) report 58% heritability for aggression among 19-month-olds; Van den Oord et al. (1996) report 69% heritability for aggression among 3-year-olds; Van der Valk et al.

(1998) report 50% heritability for externalising behaviour problems among 2–3-year-old boys and 75% for girls, and more recently the same group (van der Valk et al., 2001) report 47% heritability for externalising behaviour among 3-year-olds. However, our finding of 82% heritability for a multiple indicator latent factor of pervasive childhood antisocial behaviour is higher than the estimates, including our own, that rely on single measures of antisocial behaviour.

How is the common factor to be interpreted? Strongest genetic influence was detected specifically by using multi-informant data. For many years, clinicians and researchers have advocated collecting multiple sources of data to assess disruptive behaviour (Achenbach et al., 1987; van der Ende, 1999; Hewitt et al., 1997). There is also a strong tradition of using multiple measures in statistical latent factor models to improve measurement of disorders in research (Bank et al., 1990; Patterson, Capaldi, & Bank, 1991; Fergusson & Horwood, 1993). Each different informant contributes to increase the validity of the assessment by bringing his/her unique perspective on children's behaviour. However, these unique perspectives are also translated into low correlations across different informants. These low correlations are to be expected as they represent differences in characteristics of informants (e.g., teacher's experience, mother's insight, children's age), differences in the amount of time each informant has to observe the children (e.g., 1–2 hours for examiners vs. a school term for teachers), differences in types of setting (e.g., mothers at home vs. teachers at school) and differences in the informants' assessment skills (e.g., highly educated teachers vs. children with limited understanding of the questions). Despite normally low cross-informant correlations, our four measures of antisocial behaviour were correlated non-trivially. They tap variation common to each of the four measures, as shown by the factor analysis and loadings on the common factor from the psychometric model. This common factor represents antisocial behaviour that was observed by different informants, and is thus by definition pervasive across settings. Research has shown that such pervasive behaviour that is corroborated by many reporters represents the severest extreme of the population's symptom distribution (Baillargeon et al., 2001) and it is known to be profoundly consequential for adolescent and adult adjustment (Moffitt, 1993). Findings related to this common factor must not be discounted by the relatively small factor loadings on each of the four measures. The common factor accounts only for a significant small proportion of the variance in any given measure of antisocial behaviour, because pervasive antisocial behaviour is rare in the population. Our study suggests that in addition to being pervasive, extreme, predictive, and rare, this behaviour is also highly heritable.

How are the source-specific factors interpreted? The genetic contribution we observed for our multi-informant factor of pervasive antisocial behaviour (82% of the common variation) was higher than the heritability estimates we showed for situational antisocial behaviour (34% of the unique variation of children's report, 56% for examiner-observers, 65% for mothers, and 71% for teachers). Situational antisocial behaviours, which account for most variance in each of the four measures, may represent a more prevalent and less severe form of antisocial behaviour. These somewhat smaller genetic effects we observed may indicate that situational antisocial behaviour is less heritable than pervasive antisocial behaviour.

Our finding of 82% heritability for pervasive childhood antisocial behaviour also contrasts against the estimate of 40% heritability for adolescent and adult antisocial behaviour from a recent meta-analysis of 51 studies (Rhee & Waldman, 2002). Some research supports our high heritability estimate by suggesting that genetic etiological processes contribute more to the form of antisocial behaviour that is pervasive across settings and begins in childhood, than to the more prevalent situational form that emerges in adolescence. A Minnesota twin study found early-onset antisocial behaviour to be strongly familial and substantially heritable, in contrast to adolescent-onset which was less familial and largely influenced by environmental factors (Taylor, Iacono, & McGue, 2000). Related evidence comes from behaviour genetic analyses of the Aggression and Delinquency scales from the CBCL. The CBCL Aggression scale is thought to be associated with the early-onset life-course persistent trajectory of antisocial development because its mean scores are stable from early childhood forward, and because it measures antisocial personality traits and physical aggression. In contrast, the CBCL Delinquency scale is associated with the adolescent-onset-transient type, because its mean scores rise during adolescence and it measures rule-breaking behaviour (Stranger, Achenbach, & Verhulst, 1997). Twin and adoption studies of these scales report higher heritability for Aggression (around 60%) than Delinquency (around 30–40%) (Deater-Deckard & Plomin, 1999; Edelbrock, Rende, Plomin, & Thompson, 1995; Eley, Lichtenstein, & Stevenson, 1998). A recent longitudinal study of 1000 Swedish twins (Eley, Lichtenstein, & Moffitt, in press) reported that continuity from childhood (age 8–9 years) to adolescence (age 13–14 years) in the CBCL Aggression scale was largely mediated by genetic influences, suggesting that childhood-onset antisocial behaviour is a more stable heritable phenotype than the adolescent-onset form of antisocial behaviour.

These findings, taken together with high heritability estimates from the four aforementioned large studies of twins aged 5 years and under, indicate that childhood-onset antisocial behaviour may be much more heritable than was previously deduced based on the modest heritability estimates obtained in studies of adolescents and adults. If this is true, it recommends that research and theory on the etiology of childhood antisocial behaviour must look beyond the current focus on socioeconomic contexts and parenting processes, to incorporate genetic explanations and develop new theories of nature–nurture interplay (Hill, 2002). In particular, children with pervasive antisocial behaviour problems may be more appropriate samples for molecular genetic research than samples of adolescent or adult offenders.

This study also has implications for measurement of behaviour problems. In the past, researchers have dismissed some forms of ratings; for example, mothers were believed to be biased, children were assumed to be incapable of self-reporting on their behaviour, and observational methods needed improvements in construct validity (Deater-Deckard, 2000; Leve, Winebarger, Fagot, Reid, & Goldsmith, 1998). The present study indicates that these reservations are unfounded. First, based on previous research, we expected moderate correlations between the four informants' reports of antisocial behaviour, but despite the modesty of these correlations, we also found significant shared variation in the informants' reports, indicating that different informants were reporting on the same construct. Second, we observed comparable estimates of genetic and environmental components across the four informants, suggesting that antisocial behaviour among young children has similar genetic variation no matter who reports it. Third, beyond the information that was agreed-upon by the informants, each one of them reported unique information that we could deduce was meaningful because it was accounted for by genetic factors, as opposed to measurement error. The findings suggest that researchers studying children's behaviour disorders should try to collect data from multiple, different sources (Loeber, Green, Lahey, & Stouthamer-Loeber, 1990; Jensen et al., 1995; van der Ende, 1999). Also, our results with respect to the unique information provided by each informant indicate that using all the information from each rater to create composite scores will capture more meaningful variation than restricting composites to only the information agreed upon by all raters. This recommends the strategy of counting symptoms if they are reported by any source, which has been suggested previously by others (Bird, Gould, & Staghezza, 1992; Piacentini, Cohen, & Cohen, 1992; Jensen et al., 1999).

Finally, this study has implications for clinical intervention. High heritability estimates found for antisocial behaviour among young children imply that clinicians treating children can expect to encounter antisocial behaviour in other family members, particularly in their young patients' parents.

Research shows that parents' antisocial behaviour promotes children's maladjustment via genetic, but also strong environmental rearing effects (Jaffee, Moffitt, Caspi, & Taylor, 2003). Previous studies also show that parents' antisocial behaviour is an important factor contributing to non-compliance and drop-out from therapeutic programs (Kazdin & Mazurick, 1994). Preventionists who translate research into therapeutic prevention programmes must develop strategies for dealing not only with difficult youngsters, but also unskilled and uncooperative parents.

Acknowledgements

Correspondence to

Louise Arseneault, Box Number P080, Institute of Psychiatry, De Crespigny Park, London SE5 8AF, UK; Email: l.arseneault@iop.kcl.ac.uk

References

Ablow, J.C., & Measelle, J.R. (1999). The Berkeley Puppet Interview (BPI): Interviewing and coding system manuals. Eugene, OR: University of Oregon, Department of Psychology.

Ablow, J.C., Measelle, J.R., Kramer, H.C., Harrington, R., Luby, J., Smider, N., Dierker, L., Clark, V., Dubicka, B., Heffelfinger, A., Essex, M.J., & Kupfer, D.J. (1999). The MacArthur Three-City Outcome Study: Evaluating multi-informant measures of young children's symptomatology. *Journal of the American Academy of Child and Adolescent Psychiatry, 38*, 1580–1590.

Achenbach, T.M. (1991a). *Manual for the Child Behavior Checklist/4–18 and 1991 Profile.* Burlington, VT: University of Vermont, Department of Psychiatry.

Achenbach, T.M. (1991b). *Manual for the Teacher's Report Form and 1991 Profile.* Burlington, VT: University of Vermont, Department of Psychiatry.

Achenbach, T.M., McConaughy, S.H., & Howell, C.T. (1987). Child/adolescent behavioral and emotional problems: Implication of cross-informant correlations

for situational specificity. *Psychological Bulletin, 101*, 213–232.

American Psychiatric Association (1994). *Diagnostic and Statistical Manual of Mental Disorders* (4th edn). Washington, DC: American Psychiatric Association.

Baillargeon, R.H., Boulerice, B., Tremblay, R.E., Zoccolillo, M., Vitaro, F., & Kohen, D.E. (2001). Modeling interinformant agreement in the absence of a 'gold standard'. *Journal of Child Psychology and Psychiatry, 42*, 463–473.

Bank, L., Dishion, T.J., Skinner, M.L., & Patterson, G.R. (1990). Method variance in structural equation modeling: Living with 'glop'. In G.R. Patterson (Ed.), *Depression and aggression in family interaction* (pp. 247–279). Hillsdale, NJ: Lawrence Erlbaum Associates.

Bird, H.R., Gould, M.S., & Staghezza, B. (1992). Aggregating data from multiple informants in child psychiatry epidemiological research. *Journal of the American Academy of Child and Adolescent Psychiatry, 31*, 78–85.

Birth Statistics. (1996). *Series FM1 no. 23.* London: HMSO.

Bollen, K.A. (1989). *Structural equations with latent variables.* New York: Wiley.

Boyle, M.H., Offord, D.R., Racine, Y., Sanford, D., Szatmari, P., Fleming, J.E., & Price-Munn, N. (1993). Evaluation of the Diagnostic Interview for Children and Adolescents for use in general population samples. *Journal of Abnormal Child Psychology, 21*, 663–681.

Burnham, K.P., & Anderson, D.R. (1998). *Model selection and inference: A practical information-theoretic approach.* New York: Springer-Verlag.

Caspi, A. (2000). The child is father of the man: Personality continuities from childhood to adulthood. *Journal of Personality and Social Psychology, 78*, 158–172.

Caspi, A., Henry, B., McGee, R.O., Moffitt, T.E., & Silva, P.A. (1995). Temperamental origins of child and adolescent behavior problems: From age 3 to age 15. *Child Development, 66*, 55–68.

Caspi, A., Moffitt, T.E., Newman, D.L., & Silva, P.A. (1996). Behavioral observations at age 3 predict adult psychiatric disorders: Longitudinal evidence from a birth cohort. *Archives of General Psychiatry, 53*, 1033–1039.

Chilcoat, H.D., & Bresleau, N. (1997). Does psychiatric history bias mothers' reports? An application of a new analytic approach. *Journal of the American Academy of Child and Adolescent Psychiatry, 36*, 971–979.

Deater-Deckard, K. (2000). Parenting and child behavioral adjustment in early childhood: A quantitative genetic approach to studying family process. *Child Development, 71*, 468–484.

Deater-Deckard, K., & Plomin, R. (1999). An adoption study of the etiology of teacher and parent reports of externalising behavior problems in middle childhood. *Child Development, 70*, 144–154.

DiLalla, L.F. (2002). Behavior genetics of aggression in children: Review and future directions. *Developmental Review, 22*, 593–622.

DiLalla, L.F., & Gottesman, I.I. (1989). Heterogeneity of causes for delinquency and criminality: Lifespan

perspectives. *Development and Psychopathology*, *1*, 339–349.

Dionne, G., Tremblay, R.E., Boivin, M., Laplante, D., & Pérusse, D. (2003). Physical aggression and expressive vocabulary in 19-month-old twins. *Developmental Psychology*, *37*, 261–273.

Dumas, J.E., & Wekerle, C. (1995). Maternal reports of child behavior problems and personal distress as predictors of dysfunctional parenting. *Development and Psychopathology*, *7*, 465–479.

Edelbrock, C., Costello, A.J., Dulcan, M.K., Kalas, R., & Conover, N.C. (1985). Age differences in the reliability of the psychiatric interview of the child. *Child Development*, *56*, 265–275.

Edelbrock, C., Rende, R., Plomin, R., & Thompson, L.A. (1995). A twin study of competence and problem behavior in childhood and early adolescence. *Journal of Child Psychology and Psychiatry*, *36*, 775–785.

Eley, T.C., Lichtenstein, P., & Moffitt, T.E. (in press). A longitudinal behavioral genetic analysis of the etiology of aggressive and non-aggressive antisocial behavior. *Development and Psychopathology*.

Eley, T.C., Lichtenstein, P., & Stevenson, J. (1998). Sex differences in the etiology of aggressive and non-aggressive antisocial behavior: Results from two twin studies. *Child Development*, *70*, 155–168.

Farrington, D.P., Gallagher, B., Morley, L., St. Ledger, R.J., & West, D. (1988). Are there any successful men from criminogenic backgrounds? *Psychiatry*, *51*, 116–130.

Fergusson, D.M., & Horwood, L.J. (1993). The structure, stability and correlations of the trait components of conduct disorder, attention deficit and anxiety/withdrawal reports. *Journal of Child Psychology and Psychiatry*, *34*, 749–766.

Fergusson, D.M., Lynskey, M.T., & Horwood, L.J. (1993). The effect of maternal depression on maternal ratings of child behavior. *Journal of Abnormal Child Psychology*, *21*, 245–269.

Gjone, H., & Novik, T.S. (1995). Parental ratings of behavioural problems: A twin and general population comparison. *Journal of Child Psychology and Psychiatry*, *36*, 1213–1224.

Goldsmith, H.H., & Gottesman, I.I. (1981). Origins of variation in behavioural style: A longitudinal study of temperament in young twins. *Child Development*, *52*, 91–103.

Haapasalo, J., & Tremblay, R.E. (1994). Physically aggressive boys from age 6 to 12: Family background, parenting behavior, and prediction of delinquency. *Journal of Consulting and Clinical Psychology*, *62*, 1044–1052.

Hewitt, J.K., Silberg, J.L., Neale, M.C., Eaves, L., & Erickson, M. (1992). The analysis of parental ratings of children's behavior using LISREL. *Behavior Genetics*, *22*, 293–317.

Hewitt, J.K., Silberg, J.L., Rutter, M., Simonoff, E., Meyer, J.M., Maes, H.H., Pickles, A., Neale, M.C., Loeber, R., Erickson, M.T., Kendler, K.S., Heath, A.C., Truett, K.R., Reynolds, C.A., & Eaves, L.J. (1997). Genetics and developmental psychopathology: 1. Phenotypic assessment in the Virginia Twin Study of Adolescent Behavioral Development. *Journal of Child Psychology and Psychiatry*, *38*, 943–963.

Hill, J. (2002). Biological, psychological, and social processes in the conduct disorders. *Journal of Child Psychology and Psychiatry*, *43*, 133–164.

Hu, L., & Bentler, P.M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*, 1–55.

Jaffee, S.R., Moffitt, T.E., Caspi, A., & Taylor, A. (2003). Life with (without) father: The benefits of living with two biological parents depend on the father's antisocial behavior. *Child Development*, *74*, 109–126.

Jensen, P.S., Roper, M., Fisher, P., Piacentini, J.C., Canino, G., Richters, J., Rubio-Stipec, M., Dulcan, M., Goodman, S., Davies, M., Rae, D., Shaffer, D., Bird, H., Lahey, B., & Schwab-Stone, M. (1995). Test-retest reliability of the Diagnostic Interview Schedule for Children (DISC 2.1): Parent, child, and combined algorithms. *Archives of General Psychiatry*, *52*, 61–71.

Jensen, P.S., Rubio-Stipec, M., Canino, G., Bird, H., Dulcan, M., Schwab-Stone, M., & Lahey, B. (1999). Parent and child contributions to diagnosis of mental disorder: Are both informants always necessary? *Journal of the American Academy of Child and Adolescent Psychiatry*, *38*, 1569–1579.

Johnson, W., Krueger, R.F., Bouchard, T.J., & McGue, M. (2002). The personalities of twins: Just ordinary folks. *Twin Research*, *5*, 125–131.

Kagan, J. (1998). Biology and the child. In W. Damon & N. Eisenberg (Eds.), *Handbook of child psychology: Vol. 3. Social, emotional, and personality development* (pp. 177–236). New York, NY: Cambridge University Press.

Kazdin, A.E., & Mazurick, J.L. (1994). Dropping out of child psychotherapy. *Journal of Consulting and Clinical Psychology*, *62*, 1069–1074.

Keenan, K. (2001). Uncovering preschool precursors to problem behavior. In R. Loeber & D.P. Farrington (Eds.), *Child delinquents: Development, intervention, and service needs* (pp. 117–134). Thousand Oaks, CA: Sage.

Keenan, K., & Wakschlag, L.S. (2002). Can a valid diagnosis of disruptive behavior disorder be made in preschool children? *American Journal of Psychiatry*, *159*, 351–358.

Kendler, K.S. (1993). Twin studies of psychiatric illness: Current status and future directions. *Archives of General Psychiatry*, *50*, 905–915.

Kendler, K.S., Martin, N.G., Heath, A.C., & Eaves, L. (1995). Self-report psychiatric symptoms in twins and their nontwin relatives: Are twins different? *American Journal of Medical Genetics*, *60*, 588–591.

Krueger, R., Caspi, A., Moffitt, T.E., Bleske, A., & Silva, P.A. (1998). Assortative mating for antisocial behaviour: Developmental and methodological implications. *Behavior Genetics*, *28*, 173–186.

Kuntsi, J., & Eley, T.C. (unpublished work). A shared genetic etiology of hyperactive behaviour and cognitive test performance.

Leve, L.D., Winebarger, A.A., Fagot, B.I., Reid, J.B., & Goldsmith, H.H. (1998). Environmental and genetic variance in children's observed and reported maladaptive behavior. *Child Development*, *69*, 1286–1298.

Levy, F., Hay, D., McLaughlin, M., Wood, C., & Waldman, I. (1996). Twin–sibling differences in parental

reports of ADHD, speech, reading, and behavioural problems. *Journal of Child Psychology and Psychiatry*, *37*, 569–578.

Loeber, R., Green, S., Lahey, B., & Stouthamer-Loeber, M. (1990). Optimal informants on childhood disruptive behaviors. *Development and Psychopathology*, *1*, 317–337.

MacCallum, R.C., Browne, M.W., & Sugawara, H.M. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychological Methods*, *1*, 130–149.

Maynard, R.A. (1997). *Kids having kids: Economic costs and social consequences of teen pregnancy*. Washington, DC: Urban Institute Press.

McGue, M., Bouchard, T.J., Iacono, W.G., & Lykken, D.T. (1993). Behavioral genetics of cognitive ability: A life-span perspective. In R. Plomin & G.E. McClearn (Eds.), *Nature, nurture, and psychology* (pp. 59–76). Washington, DC: American Psychological Association.

McGuffin, P., Riley, B., & Plomin, R. (2001). Toward behavioral genomics. *Science*, *291*, 1232–1234.

Measelle, J.R., Ablow, J.C., Cowan, P.A., & Cowan, C.P. (1998). Assessing young children's views of their academic, social, and emotional lives: An evaluation of the self-perception scales of the Berkeley Puppet Interview. *Child Development*, *69*, 1556–1576.

Moffitt, T.E. (1993). 'Life-course-persistent' and 'adolescence-limited' antisocial behavior: A developmental taxonomy. *Psychological Review*, *100*, 674–701.

Moffitt, T.E., Caspi, A., Harrington, H., & Milne, B.J. (2002). Males on the life-course persistent and adolescence-limited antisocial pathways: Follow-up at age 26. *Development and Psychopathology*, *14*, 179–206.

Moffitt, T.E., Caspi, A., Rutter, M., & Silva, P.A. (2001). *Sex differences in antisocial behaviour: Conduct disorder, delinquency, and violence in the Dunedin Longitudinal Study*. Cambridge: Cambridge University Press.

Moffitt, T.E., & the E-Risk Study Team (2002). Teenaged mothers in contemporary Britain. *Journal of Child Psychology and Psychiatry*, *43*, 727–742.

Moilanen, I., Linna, S.-L., Ebeling, H., Kumpulainen, K., Tamminen, T., Piha, J., & Almqvist, F. (1999). Are twins' behavioural/emotional problems different from singletons'? *European Child and Adolescent Psychiatry*, *8* (Suppl. 4), 62–67.

Muthén, L.K., & Muthén, B.O. (1998). *Mplus: The comprehensive modeling program for applied researchers*. Los Angeles: Author.

Neale, M.C. (1997). *Mx: Statistical modelling* (4th edn). Richmond, VA: Medical College of Virginia, Department of Psychiatry.

Neale, M.C., & Cardon, L.R. (1992). *Methodology for genetic studies of twins and families*. Dordrecht: Kluwer Academic.

Patterson, G.R., Capaldi, D., & Bank, L. (1991). An early starter model for predicting delinquency. In D. Pepler & K. Rubin (Eds.), *The development and treatment of childhood aggression* (pp. 139–168). Hillsdale, NJ: Erlbaum.

Patterson, G.R., & Yoerger, K. (1997). A developmental model for late-onset delinquency. In D.W. Osgood (Ed.), *Motivation and delinquency* (pp. 119–177). Lincoln, NE: University of Nebraska Press.

Piacentini, J.C., Cohen, P., & Cohen, J. (1992). Combining discrepant diagnostic information from multiple sources: Are complex algorithms better than simple ones? *Journal of Abnormal Child Psychology*, *20*, 51–63.

Plomin, R., DeFries, J.C., McClearn, G.E., & McGuffin, P. (2001). *Behavioral genetics* (4th edn). New York, NY: Worth.

Plomin, R., Fulker, D.W., Corley, R., & DeFries, J.C. (1997). Nature, nurture, and cognitive development from 1 to 16 years: A parent–offspring adoption study. *Psychological Science*, *8*, 442–447.

Raftery, A.E. (1995). Bayesian model selection in social research. In P.V. Marsden (Ed.), *Sociological methodology* (pp. 111–196). Oxford, UK: Blackwell.

Rhee, S., & Waldman, I.D. (2002). Genetic and environmental influences on antisocial behavior: A meta-analysis of twin and adoption studies. *Psychological Bulletin*, *29*, 490–529.

Robins, L.N. (1966). *Deviant children grown up*. Baltimore, MD: Williams and Wilkins.

Rowe, D.C., & Osgood, D.W. (1984). Heredity and sociological theories of delinquency: A reconstruction. *American Sociological Review*, *49*, 526–540.

Rutter, M., & Silberg, J. (2002). Gene–environment interplay in relation to emotional and behavioral disturbance. *Annual Review of Psychology*, *53*, 463–490.

Rutter, M., Thorpe, K., Greenwood, R., Northstone, K., & Golding, J. (2003). Twins as a natural experiment to study the causes of mild language delay: I: Design; twin–singleton differences in language, and obstetric risks. *Journal of Child Psychology and Psychiatry*, *44*, 326–341.

Scarr, S., & McCartney, K. (1983). How people make their own environments: A theory of genotype to environment effects. *Child Development*, *54*, 424–435.

Schmitz, S., Fulker, D.W., & Mrazek, D.A. (1995). Problem behavior in early and middle childhood: An initial behavior genetic analysis. *Journal of Child Psychology and Psychiatry*, *36*, 1443–1458.

Schwab-Stone, M., Fallon, T., Briggs, M., & Crowther, B. (1994). Reliability of diagnostic reporting for children aged 6–11 years: A test-retest study of the Diagnostic Interview Schedule for Children–Revised. *American Journal of Psychiatry*, *151*, 1048–1054.

Shaw, D.S., Owens, E.B., Vondra, J.I., Keenan, K., & Winslow, E.B. (1996). Early risk factors and pathways in the development of early disruptive behavior problems. *Development and Psychopathology*, *8*, 679–699.

Simonoff, E., Pickles, A., Hervas, A., Silberg, J.L., Rutter, M., & Eaves, L. (1998). Genetic influences on childhood hyperactivity: Contrast effects imply parental rating bias, not sibling interaction. *Psychological Medicine*, *28*, 825–837.

Simonoff, E., Pickles, A., Hewitt, J., Silberg, J., Rutter, M., Loeber, R., Meyer, J., Neale, M., & Eaves, L. (1995). Multiple raters of disruptive child behavior: Using a genetic strategy to examine shared views and bias. *Behavior Genetics*, *25*, 311–326.

Simonoff, E., Pickles, A., Meyer, J., Silberg, J.L., Maes, H.H., Loeber, R., Rutter, M., Hewitt, J.K., & Eaves, L. (1997). The Virginia Twin Study of Adolescent Behavioral Development. *Archives of General Psychiatry*, *54*, 801–808.

Stevenson, J., & Goodman, R. (2001). Association between behaviour at age 3 years and adult criminality. *British Journal of Psychiatry*, *179*, 197–202.

Stranger, C., Achenbach, T.M., & Verhulst, F.C. (1997). Accelerated longitudinal comparisons of aggressive versus delinquent syndromes. *Development and Psychopathology*, *9*, 43–58.

Taylor, J., Iacono, W.G., & McGue, M. (2000). Evidence for a genetic etiology of early-onset delinquency. *Journal of Abnormal Psychology*, *109*, 634–643.

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

Tremblay, R.E., Pihl, R.O., Vitaro, F., & Dobkin, P.L. (1994). Predicting early onset of male antisocial behavior from preschool behavior. *Archives of General Psychiatry*, *51*, 732–739.

Trouton, A., Spinath, F.M., & Plomin, R. (2002). Twins Early Development Study (TEDS): A multivariate, longitudinal genetic investigation of language, cognition and behaviour problems in childhood. *Twin Research*, *5*, 444–448.

van den Oord, E.J.C.G., Boomsma, D.I., & Verhulst, F.C. (2000). A study of genetic and environmental effects on the co-occurrence of problem behaviors in three-year-old twins. *Journal of Abnormal Psychology*, *109*, 360–372.

van den Oord, E.J.C.G., Koot, H.M., Boomsma, D.I., Verhulst, F.C., & Orleveke, J.F. (1995). A twin–singleton comparison of problem behaviour in 2–3-year-olds. *Journal of Child Psychology and Psychiatry*, *36*, 449–458.

van den Oord, E.J.C.G., Verhulst, F.C., & Boomsma, D.I. (1996). A genetic study of maternal and paternal ratings of problem behaviors in 3-year-old twins. *Journal of Abnormal Psychology*, *105*, 349–357.

van der Ende, J. (1999). Multiple informants: Multiple views. In H.M. Koot, A.A.M. Crijnen, & R.F. Ferdinand (Eds.), *Child psychiatry epidemiology: Accomplishments and future directions* (pp. 39–52). Assen, The Netherlands: Van Gorcum.

van der Valk, J.C., van den Oord, E.J.C.G., Verhulst, F.C., & Boomsma, D.I. (2001). Using parental ratings to study the etiology of 3-year-old twins' problem behaviors: Different views or rater bias? *Journal of Child Psychology and Psychiatry*, *42*, 921–931.

van der Valk, J.C., Verhulst, F.C., Stroet, T.M., & Boomsma, D.I. (1998). Quantitative genetic analysis of internalising and externalising problems in a large sample of 3-year-old twins. *Twin Research*, *1*, 25–33.