Twin Research (2000) 3, 251–258

www.nature.com/tr

Parents' and teachers' ratings of problem behaviours in children: genetic and contrast effects

Jonna Kuntsi[1], Javier Gayán[2] and Jim Stevenson[3]

[1]Behavioural Sciences Unit, Institute of Child Health, University College London Medical School, London, UK
[2]Institute for Behavioral Genetics, University of Colorado, Boulder, CO, USA
[3]Centre for Research into Psychological Development, Department of Psychology, University of Southampton, UK

We obtained ratings on the Conners' scales from teachers (CTRS-28) and parents (CPRS-48) for 61 monozygotic and 64 dizygotic twin pairs, aged between 7 and 11 years. Model-fitting analyses were carried out to estimate the extent of genetic and environmental influences on problem behaviours, and to explore possible contrast effects in ratings by parents and teachers. Confirming previous findings with other measures, there was evidence of moderate to strong genetic effects on a range of problem behaviours. Parents' ratings on the Anxiety, Impulsive-Hyperactive and Learning Problem sub-scales showed significant evidence of contrast effects. There was no evidence of such rater bias or competitive sibling interaction effects in ratings by teachers, or in parents' ratings on the Conduct Problem and Psychosomatic sub-scales. Twin Research (2000) 3, 251–258.

Keywords: problem behaviours, contrast effects, CTRS-28, CPRS-48, twin study

Introduction

Genetic effects are moderately strong for a variety of problem behaviours in children and adolescents.[1] It is not only the psychiatric disorders or the 'extreme scorers' that show evidence of genetic effects, but also the symptom clusters considered as dimensions.

Such genetic data on dimensional measures of problem behaviours in general population samples of children in the mid to late childhood age range have been reported for the Child Behaviour Checklist (CBCL). The CBCL is one of the most widely used questionnaires to obtain behavioural ratings from parents. Edelbrock and colleagues[2] obtained ratings for 99 monozygotic (MZ) and 82 dizygotic (DZ) twin pairs aged between 7 and 15 years. A twin study by Schmitz and colleagues[3] included data on CBCL for 66 MZ and 137 DZ twin pairs aged between 4 and 18 years.

Overall, the two studies obtained a similar pattern of results, with the heritability estimates for the eight behavioural problem scales of the CBCL ranging from 34% to 79%. Genetic factors accounted for a substantial proportion of variance in attention problems, aggressive behaviours, somatic complaints and social problems. Slightly lower heritability estimates were reported for symptoms of anxiety and depression. For the delinquent behaviour sub-scale the findings were discrepant: Edelbrock et al[2] reported a heritability estimate of 35%, whereas Schmitz et al[3] obtained an estimate of 79%. The older average age of the children in the Edelbrock et al study provides a possible explanation for this discrepancy: delinquent behaviours shown during adolescence may be influenced less by genetic factors than antisocial behaviours among younger children.[4]

Other studies using alternative measures similarly report widespread influence of genetic factors on problem behaviours in children and adolescents. For example, the Virginia Twin Study of Adolescent Behavioural Development includes questionnaire- and interview-based data on a large, representative sample of 8–16-year-old twins.[5] Analyses on these data showed that genetic factors accounted for more than half of the variance for several different types of problem behaviours. The environmental influences accounting for the remaining variance were mostly of the non-shared kind, that is, those environmental influences which make members of the same family different from one another.

The estimates for the relative contribution of genes and environment are undoubtedly also influenced by measure- and rater-specific effects. The evidence

Correspondence: Jonna Kuntsi PhD, Behavioural Sciences Unit, Institute of Child Health, University College London Medical School, 30 Guilford Street, London WC1N 1EH, UK.
Tel: 020 7831 0975; Fax: 020 7831 7050;
E-mail: j.kuntsi@ich.ucl.ac.uk.
Received 25 May 1999; revised 1 June 2000; accepted 3 July 2000

of low or negative DZ correlations and greater variances in DZ than MZ twins in parents' ratings on certain behaviours is suggestive of contrast effects. Contrast effects refer to the negative influence of the phenotype or behaviour of one individual on that of another and may reflect either true phenotypic effects or rater bias. True competitive sibling interaction means that the more one twin shows the behaviour, the less does the other. The rater bias explanation suggests that the more one twin is perceived to show the behaviour, the less extreme the perception of the other twin's behaviour. Dominance effects (non-additive genetic effects) can also produce lowered DZ correlations, but they would not be expected to produce negative DZ correlations. In modelling analyses of twin data it is possible to remove contrast effects from the variance explained.[6]

Evidence for contrast effects in parental ratings has been obtained for attention deficit hyperactivity disorder (ADHD),[5,7,8] anxiety[5] and temperament traits,[9,10] but not for conduct problems.[5] Analyses by Simonoff and colleagues[11] suggest that for parents' ratings on hyperactivity the contrast effects reflect rater bias rather than true sibling interaction. Teacher ratings do not seem to show evidence of contrast effects.[5] However, teacher ratings are not a gold standard either, but may reflect a different bias: ratings on hyperactivity made by the same teacher were more highly correlated within twin pairs than ratings made by different teachers.[11] Teachers may vary in their expectations of behaviour, which is then reflected in their ratings ('correlated errors'), or they may have difficulty attributing behaviour to the correct child ('twin confusion').

The present study aimed to investigate this issue of possible contrast effects for a wider range of problem behaviours than has been investigated previously. A second aim was to estimate the extent of genetic effects on the various types of problem behaviours measured by the Conners' Teacher and Parent Rating Scales.[12] Although the Conners' scales are widely used to obtain ratings on a range of problem behaviours in children, we are not aware of any previous twin studies using these scales. It is important to investigate whether similar estimates of genetic and environmental effects are obtained across different rating scales that claim to measure similar types of problem behaviours. A limitation in much of the previous research on the heritability of problem behaviours and on contrast effects is the often exclusive focus on parent ratings. We obtained both teacher and parent ratings for the same sample, enabling an investigation of whether similar heritability estimates, and possibly contrast effects, are obtained across raters.

Method

Sample and procedure

The twins were recruited from a general population sample of same-sex twins aged between 7 and 11 years. We approached all primary schools (n = 2439), including special schools, in 16 Local Education Authorities in southern England. We asked the class teachers of any twins fulfilling the criteria for our study (same-sex twins; date of birth between 1 September 1985 and 1 September 1990) to complete the Conners' Teacher Rating Scale (CTRS-28),[12] separately for each twin.

We received replies from 66.8% (1629) of the schools we contacted. Of those schools which replied to our letter, only 59 indicated that they did not wish to take part in the study. In 858 of the schools there were no twins fulfilling the criteria for our study. We received the CTRS-28 for 1316 twin pairs. However, it was possible only to determine the zygosity for a sub-sample of these twins, those for whom we also obtained parent ratings.

As the last stage of the study involved assessing hyperactive and control twins on psychological tests, the results of which we report in separate publications,[13,14] at the second stage of the screening process hyperactive and control twin pairs were selected based on teacher ratings. To be included in the hyperactive group, one or both twins had to score above a cut-off point of a T-score of 64, that is 1.5 standard deviations above the mean[12] on the Hyperactivity sub-scale of the CTRS-28. To be chosen as controls, both twins had to score below the cut-off point on the Hyperactivity subscale. We asked the parents of 392 twin pairs meeting these criteria to complete the Conners' Parent Rating Scale (CPRS-48)[12] for each twin. We received the CPRS-48 from 66.4% (174) of the hyperactive families and 72.3% (94) of the control families. The parents completed the CPRS-48 on average three months after the teachers had completed the CTRS-28.

Because this sample of 268 twin pairs has an excess of hyperactive children, we created a sample representative of the general population for the analyses we report here. This representative sample was chosen as follows. Based on the teacher and parent ratings we obtained it was possible to estimate that, using our hyperactivity criterion, approximately 5% of the children in the general population would be pervasively hyperactive and 12–15% situationally hyperactive (above the cut-off point only on parent or teacher questionnaire). The twins within a pair were arbitrarily designated either Twin A or Twin B. We first chose randomly from twins A so that the proportions of situationally hyperactive, pervasively hyperactive and control children would equal approximately these proportions estimated for

https://doi.org/10.1375/twin.3.4.251 Published online by Cambridge University Press

a general population sample. We then deleted further pairs in which twin B was situationally or pervasively hyperactive, until the total proportions of situationally hyperactive, pervasively hyperactive and control children were close to the figures estimated for the general population.

There are 125 pairs in this sample: 61 MZ and 64 same-sex DZ pairs. Of these 250 children, 6% were pervasively hyperactive, 15.6% were situationally hyperactive and 78.4% were controls. Mean age was 8.0 years (SD = 1.39 years) and 44.8% of them were girls.

Zygosity determination We determined zygosity of the twins using the Twin Similarity Questionnaire (TSQ).[15] This is a short questionnaire which includes items regarding the physical similarity (hair colour, eye colour, weight, height and complexion) and physical confusability of the twins. Parents rate their twins either on a dichotomous scale (yes/no) or, for some items, on a scale from 0 to 2. The higher the score, the more similar the twins are in appearance; the maximum score is 20. The general rule we used was to classify twins who obtained a score of 13 or higher on the questionnaire as MZ and those who obtained a score of 12 or lower as DZ. Questionnaire methods have been shown to be more than 90% accurate in determining zygosity.[16]

We also took a photograph of each twin pair whom we tested. For those cases who were on the MZ/DZ borderline on the TSQ (scoring 11–14), we used the photographs to determine their zygosity: three raters independently classified the twin pairs as MZ or DZ based on the photographs. In the very rare cases when the raters disagreed about the twins' zygosity, we either rang the parents to inquire whether the twins' zygosity had been determined using blood tests or we obtained further ratings from two other raters.

Measures

The measures used in this study were the shorter version (CPRS-48 and CTRS-28)[12] of the Conners' Parent and Teacher Rating Scales. The parent scale provides the following dimensions: Conduct Problem, Learning Problem, Psychosomatic, Impulsive-Hyperactive, Anxiety and Hyperactivity Index. The dimensions obtained from the teacher scale are the following: Conduct Problem, Hyperactivity, Inattentive-Passive and Hyperactivity Index. We did not use the Hyperactivity Index in the present study, as it consists of items from the other subscales and therefore is not a separate dimension as such (it did not emerge as a factor in the original factor analysis). On both scales each item is rated as not at all present, just a little present, pretty much present, or very much present (scored 0 to 3, respectively, with higher scores indicating greater severity). See Tables 1 and 2 for the items comprising each of the dimensions. The original report[12] provided normative data, on which the T-scores are based, for the CPRS-48 and CTRS-28.

Method of analysis

The fact that DZ twins share approximately half their genes and MZ twins all their genes provides the basis for analysing twin data. Based on this genetic relatedness and the assumption that the environments for MZ and DZ twins are roughly equal (the equal environments assumption), it is possible to quantify the importance of genetic and environmental factors in causing differences between individuals on traits or abilities. Evidence from several studies supports the equal environments assumption.[17]

In the basic equation, variance in a phenotype (V_p) is divided into that due to additive genetic factors (V_a) and that due to the environment (common or shared, V_c, and non-shared V_e):

$$V_p = V_a + V_c + V_e$$

The variance due to additive genetic factors (V_a) thus represents the individual differences heritability

Table 1 Items comprising the dimensions of CTRS-28

I	Conduct Problem
	Acts 'smart' (impudent or sassy)
	Temper outbursts and unpredictable behavior
	Overly sensitive to criticism
	Pouts and sulks
	Mood changes quickly and drastically
	Quarrelsome
	Denies mistakes or blames others
	Uncooperative with teacher
II	Hyperactivity
	Restless in the 'squirmy' sense
	Makes inappropriate noises when s/he shouldn't
	Demands must be met immediately
	Disturbs other children
	Restless, always up and on the go
	Excitable, impulsive
	Excessive demands for teacher's attention
III	Inattentive-Passive
	Distractibility or attention span a problem
	Daydreams
	Appears to be easily led by other children
	Appears to lack leadership
	Fails to finish things that s/he starts
	Childish and immature
	Easily frustrated in efforts
	Difficulty in learning

https://doi.org/10.1375/twin.3.4.251 Published online by Cambridge University Press

Table 2 Items comprising the dimensions of CPRS-48

I	Conduct Problem
	Sassy to grown-ups
	Carries a chip on his/her shoulder
	Destructive
	Denies mistakes or blames others
	Quarrelsome
	Bullies others
	Fights constantly
	Basically an unhappy child
II	Learning Problem
	Difficulty in learning
	Fails to finish things
	Distractibility or attention span a problem
	Easily frustrated in efforts
III	Psychosomatic
	Headaches
	Stomach aches
	Other aches and pains
	Vomiting or nausea
IV	Impulsive-Hyperactive
	Excitable, impulsive
	Wants to run things
	Restless in the 'squirmy' sense
	Restless, always up and on the go
V	Anxiety
	Fearful (of new situations, new people or places, going to school)
	Shy
	Worries more than others (about being alone, illness or death)
	Lets self be pushed around

estimate (h^2). If the within-pair correlations between twins indicate that there may be some additive genetic effects (the MZ correlation is larger than the DZ correlation), the next step is to carry out more formal model-fitting analyses. We used the structural equation modelling programme Mx[18] to analyse the present data.

In model fitting with twin data, the MZ and DZ covariance matrices provide the data against which the model is tested. Latent variables in this type of analysis are the A, C and E terms. Figure 1 shows the simple univariate model (the full ACE model[6]). The covariance between the A terms is set to 1.0 for MZ twins and to 0.5 for DZ twins. The covariance between the C terms is set to 1.0 both for MZ and DZ twins, as this captures the shared environmental factors. By definition, there is no covariance between the non-shared environmental factors (E). The E term also contains variance due to error.

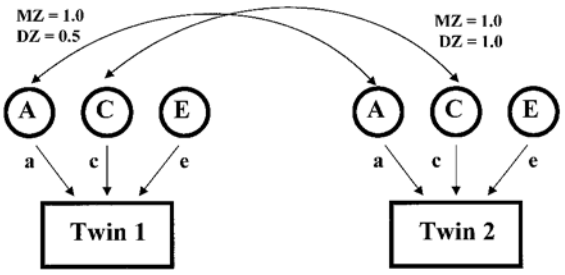

Figure 1 Univariate analysis of twin data

The full ACE model is fitted first. Then parameters which do not significantly contribute to the fit of the model are dropped. Because the E term includes measurement error, this term is not usually dropped in univariate analyses. Two models, the AE and CE models, are nested within the full model (subsets of free parameters in these models are contained in the full model). For nested models, the change in the chi-square value is used to determine which model provides the best fit for the data. For non-nested models, the AIC (Akaike's information criteria) and RMSEA (root mean square error of approximation) values are used to compare the fit of alternative models. A good fit is indicated by a negative AIC value[19] and an RMSEA value below 0.1, with an RMSEA value below 0.05 indicating very good fit.[20]

The model which includes contrast effects (ABE) is represented by reciprocal paths 'b' between the twins' phenotypes. The AE model is nested within the ABE model. If the C term is not significant, the univariate model with dominance effects (the ADE model) can be fitted to the data. The within-pair correlation for dominance is 1.0 for MZ pairs and 0.25 for DZ pairs. The AE model is nested within the ADE model and therefore the significance of the D effect can be tested by changes in chi-square. Full details of the testing of these alternative models are given in Neale and Cardon.[6]

Results

Univariate ACE models were fitted to each of the rating scale dimensions (T-scores). Following the rule of parsimony, the model with the fewest parameters, which did not significantly worsen the fit, was chosen as the best-fitting model. The AIC and RMSEA fit indices were used to compare the fit of non-nested models. In some cases the data suggested that a contrast effect model might be needed (very low or negative DZ twin correlations and greater variance in DZ than MZ pairs) and this possibility was also tested for each of the subscales. The ADE models were also fitted to the data, to enable a comparison between the contrast effect model and a model which includes genetic non-additivity. It is not possible using data from MZ and DZ twin pairs alone to estimate more complex models, such as ACDE or ACBE.[6]

Tables 3 and 4 report the fit of the models for each sub-scale of the CTRS-28 and CPRS-48, including the change in chi-square values between ACE and

https://doi.org/10.1375/twin.3.4.251 Published online by Cambridge University Press

Table 3 Fit of the models for the Teacher Conners' (CTRS-28) sub-scales

	χ^2	df	p	AIC	RMSEA	ACE vs AC/CE $\triangle\chi^2$	$\triangle$df	p
Conduct Problem								
ACE	1.92	3	0.59	–4.08	0.03			
ADE	2.62	3	0.45	–3.38	0.05			
ABE	2.08	3	0.56	–3.92	0.03			
AE	2.62	4	0.62	–5.38	0.03	0.70	1	ns
CE	7.61	4	0.11	–0.39	0.12	5.69	1	<0.05
Hyperactivity								
ACE	3.39	3	0.33	–2.61	0.06			
ADE	3.36	3	0.34	–2.64	0.06			
ABE	3.56	3	0.31	–2.44	0.06			
AE	3.39	4	0.50	–4.61	0.04	0.00	1	ns
CE	8.35	4	0.08	0.35	0.12	4.96	1	<0.05
Inattentive-Passive								
ACE	1.37	3	0.71	–4.63	0.00			
ADE	1.80	3	0.62	–4.20	0.00			
ABE	1.48	3	0.69	–4.52	0.00			
AE	1.80	4	0.77	–6.20	0.00	0.43	1	ns
CE	16.11	4	0.003	8.11	0.21	14.74	1	<0.001

AIC = Akaike's information criteria; RMSEA = Root mean square error of approximation; bold indicates a best-fitting model.

Table 4 Fit of the models for the Parent Conners' (CPRS-48) sub-scales

	χ^2	df	p	AIC	RMSEA	ACE vs AC/CE $\triangle\chi^2$	$\triangle$df	p
Conduct Problem								
ACE	6.15	3	0.11	0.15	0.11			
ADE	7.68	3	0.05	1.68	0.14			
ABE	6.41	3	0.09	0.41	0.11			
AE	7.68	4	0.10	–0.32	0.09	1.13	1	ns
CE	7.53	4	0.11	–0.47	0.09	1.38	1	ns
Learning Problem								
ACE	7.15	3	0.07	1.15	0.14			
ADE	3.47	3	0.33	–2.53	0.05			
ABE	3.03	3	0.39	–2.97	0.05			
AE	7.15	4	0.13	–0.85	0.10	0.00	1	ns
CE	15.28	4	0.004	7.28	0.21	8.13	1	<0.01
Psychosomatic								
ACE	7.99	3	0.05	1.99	0.14			
ADE	5.90	3	0.12	–0.10	0.08			
ABE	4.88	3	0.18	–1.12	0.08			
AE	7.99	4	0.09	–0.01	0.09	0.00	1	ns
CE	12.46	4	0.01	4.46	0.18	4.47	1	<0.05
Impulsive-Hyperactive								
ACE	15.75	3	0.001	9.75	0.26			
ADE	11.34	3	0.01	5.34	0.21			
ABE	7.16	3	0.07	1.16	0.15			
AE	15.75	4	0.003	7.75	0.22	0.00	1	ns
CE	21.69	4	0.001	13.69	0.26	5.94	1	<0.05
Anxiety								
ACE	29.11	3	0.001	23.11	0.36			
ADE	19.52	3	0.001	13.52	0.27			
ABE	8.89	3	0.03	2.89	0.17			
AE	29.11	4	0.001	21.11	0.30	0.00	1	ns
CE	38.18	4	0.001	30.18	0.37	9.07	1	<0.01

AIC = Akaike's information criteria; RMSEA = Root mean square error of approximation; bold indicates a best-fitting model.

https://doi.org/10.1375/twin.3.4.251 Published online by Cambridge University Press

AE/CE models (other comparisons between nested models are reported in the text). Table 5 shows the parameter estimates for the best-fitting models, including 95% confidence intervals, as well as the phenotypic twin correlations and standard deviations. The parameter estimates are the squared path coefficients representing the proportion of variance explained. For the contrast effect models (ABE), the total variance was corrected by subtracting out the extra variation created by the contrast effects.[5]

For each of the subscales, the fit of the ACE model was first compared to the fit of the AE and CE models (Tables 3 and 4). For the Conduct Problem subscale of the CTRS-28, the change in the chi-square value was non-significant from the ACE model to the AE model, but significant from the ACE model to the CE model. This indicates that the AE model provides a more parsimonious fit for the data than the ACE model and that the common environment is not making a significant contribution. The change in the chi-square values from both the ADE and ABE models to the AE model ($\Delta\chi^2$ = 0.00 and $\Delta\chi^2$ = 0.54, Δdf = 1, respectively) were also non-significant. As dropping the D or B terms did not significantly worsen the fit of the model, the AE model is chosen as the best-fitting model. The AIC and RMSEA values also indicate that the AE model provides a very good fit for the data. Heritability of teacher-rated conduct problems is estimated at 69% (Table 5).

For the Hyperactivity sub-scale of the CTRS-28, the AE model similarly provided a better fit for the data than the ACE or CE models. Comparisons between the AE model and both the ABE ($\Delta\chi^2$ = 0.17, Δdf =) and the ADE ($\Delta\chi^2$ = 0.03, Δdf = 1) models produced a non-significant chi-square result. The best-fitting AE model indicates that genetic factors account for approximately 57% of the variance in teacher-rated hyperactivity. The AE model provided the most parsimonious fit also for teacher ratings on the Inattentive-Passive subscale: the change in the chi-square value was non-significant when the AE model was compared with the ACE (Table 3), ADE ($\Delta\chi^2$ = 0.00, Δdf = 1) and ABE ($\Delta\chi^2$ = 0.32, Δdf = 1) models. The data suggest a high heritability of 80%. For each of the dimensions, the non-shared environment (and measurement error) account for the remaining variance.

Parent ratings on the CPRS-48 produced a different pattern of results. For the Conduct Problem subscale, neither the AE model nor the CE model significantly worsened the fit, as compared with the ACE model (Table 4). However, the fit of the AE and CE models were equally good and therefore one could not choose between these two models. As the ADE and ABE models provided a fit that is less good than that of the ACE model, the most conservative option here was to choose the ACE model, but this calls for caution in interpreting the results. The ACE model produced a rather low heritability estimate of 29%, with a 95% confidence interval ranging from 0% to 69% (Table 5). The AE model would have produced a higher heritability estimate of 59% (95% confidence interval 42–71%), a figure not far from the heritability estimate based on parent ratings on conduct problems.

For each of the CPRS-48 subscales of Learning Problem, Impulsive-Hyperactive and Anxiety, the contrast effect model (ABE) provided the best fit for the data. In each case, the change in chi-square from the ACE model to the AE model was non-significant, but was significant from the ACE to the CE model (Table 4). The fit of the AE model was significantly worse than the fit of the ABE model, as indicated by a significant chi-square value (Learning Problem $\Delta\chi^2$ = 4.12, Δdf = 1, $P < 0.5$; Impulsive-Hyperactive $\Delta\chi^2$ = 8.59, Δdf = 1, $P < .01$; Anxiety $\Delta\chi^2$ = 20.22, Δdf = 1, $P < .001$). Further, for each sub-scale the fit indices indicated that the ABE model fitted the data better than the ADE model (Table 4). Based on the ABE models, genetic factors account for 72% of variance in parent ratings on the Learning Problem

Table 5 Contribution of additive genetic (h^2), shared environmental (c^2) and non-shared environmental (e^2) components to total variation in problem behaviours (estimated from best-fitting models), contrast effects, twin correlations and standard deviations

	Component: % variance (95% confidence interval)			Contrast ($\times$ 100)	MZ			DZ		
Sub-scale	h^2	c^2	e^2		r	SD_1	SD_2	r	SD_1	SD_2
Teacher Conners'										
Conduct Problem	69 (56–79)	–	31 (21–44)	–	0.67	12.10	11.75	0.44	11.43	13.25
Hyperactivity	57 (38–70)	–	43 (30–62)	–	0.57	9.86	11.72	0.27	10.58	11.68
Inattentive-Passive	80 (70–86)	–	20 (14–30)	–	0.79	9.66	10.49	0.47	10.23	10.69
Parent Conners'										
Conduct Problem	29 (0–69)	27 (0–58)	44 (30–63)	–	0.56	10.31	13.34	0.43	12.20	12.89
Learning Problem	72 (44–84)	–	28 (16–56)	–16 (–30/–1)	0.57	13.89	15.96	–0.02	15.05	14.25
Psychosomatic	38 (15–57)	–	62 (43–85)	–	0.44	12.44	10.97	–0.02	13.87	10.87
Impulsive-Hyperactive	71 (47–83)	–	29 (17–53)	–24 (–26/–10)	0.48	8.71	10.57	–0.01	10.85	12.94
Anxiety	85 (73–92)	–	15 (8–27)	–31 (–41/–19)	0.62	10.94	9.24	–0.26	12.53	9.59

https://doi.org/10.1375/twin.3.4.251 Published online by Cambridge University Press

sub-scale, 71% of variance on the Impulsive-Hyperactive sub-scale and 85% of the variance on the Anxiety sub-scale.

For the Psychosomatic subscale of the CPRS-48, the AE model provided the best fit for the data, with heritability estimated at 38%. Comparisons between the AE and each of the nested models produced a non-significant chi-square result (ACE Table4; ADE $\Delta^2_\chi = 2.09$ $\Delta df = 1$, ABE $\Delta^2_\chi = 3.11$, $\Delta df = 1$).

Discussion

Whereas several previous twin studies have estimated the extent to which genetic and environmental factors influence individual differences in scores on behavioural rating scales, to our knowledge this is the first twin study to report such data for the Conners' scales.

Teacher ratings on the Inattentive-Passive dimension suggested high heritability: approximately 80% of the variance was due to genetic effects. This accords with the literature showing strong genetic effects on attention problems[2,3,21] Not all the items of this sub-scale focus on inattentiveness, however. The other items include, for example, 'difficulty in learning', 'appears to be easily led by other children' and 'childish and immature'. The findings of rather high heritabilities for the hyperactivity dimensions similarly confirm the general pattern of findings from previous studies on hyperactivity or ADHD (a more extensive discussion of this is available[14]).

The present data also confirm the finding of genetic factors influencing somatic symptoms, although the heritability estimate was somewhat lower than the heritability estimates from the CBCL studies.[2,3] For parents' ratings on anxiety, our findings suggest higher heritability (85%) than the CBCL studies (34% and 50%). The two scales are not directly comparable, however, as the CBCL sub-scale includes both anxiety and depression items. It is important also to remember that all these figures are indeed estimates, with associated confidence intervals. Studies using other parent questionnaires suggest that genetic factors account for approximately 60% of the variance in anxiety symptoms in children and adolescents.[5,22] Genetic factors explained a substantial amount of variance on the Learning Problem dimension too. This sub-scale includes items such as 'difficulty in learning' and 'fails to finish things'.

The heritability of conduct problems was estimated at 69% based on teacher report and at 29% based on parent report. Caution is needed when interpreting the parent-report data, however, as it was difficult to choose between the various models. The heritability estimate based on teacher ratings on conduct problems is in line with previous reports of the extent to which genetic factors account for aggressive behaviour.[2,3] The Conners' scales do not have separate sub-scales for aggressive behaviour and delinquent behaviour, but the emphasis on the Conduct Problem sub-scale is on aggressive behaviours.

A specific aim of this study was to examine the extent to which contrast effects may influence ratings by parents or teachers. We obtained evidence of such rater bias or competitive sibling interaction effects on parents' ratings on the Anxiety, Impulsive-Hyperactive and Learning Problem subscales. These findings extend previous findings which had indicated contrast effects on parental ratings on anxiety[5] and hyperactivity or ADHD[5,7,8]

Confirming previous findings,[5] there was no evidence of contrast effects in ratings by teachers. Parents' ratings on the Conduct Problem and Psychosomatic subscales also showed no evidence of contrast effects. Simonoff and colleagues[11] have suggested that norms may be more clear cut for conduct problems, which may explain the lack of rater bias in parents' ratings on such behaviours.

Previous research has shown that the contrast effects in parents' ratings on hyperactivity are likely to reflect rater bias rather than true phenotypic effects.[11] The present study design did not enable us to distinguish between these two possibilities. Future studies could explore this issue further in relation to the other behavioural ratings which show contrast effects.

Although the emphasis in this paper is on the findings of genetic effects, the results also provide evidence of significant environmental effects on problem behaviours. This issue of environmental effects relates to a limitation of this study: the modest sample size results in limited power to detect shared environmental effects in particular. Nonetheless, the lack of shared environmental effects on most behavioural dimensions and disorders is a strong finding across studies.[17] The environmental factors which influence most behaviours seem not to be of the kind which are shared between members of the same family but rather to be those factors which are specific to each individual. On the other hand, Simonoff and colleagues[11] point out that it is difficult to detect both shared environmental effects and contrast effects, if both were present; this requires large sample sizes and either extended genetic designs or multiple informants.

Another limitation of the present study relates to the nature of the sample. Because of the excess of hyperactive children in the larger sample of 268 twin pairs, we had to create a sample representative of the general population by excluding over half of the twin pairs from the analyses. However, we showed

https://doi.org/10.1375/twin.3.4.251 Published online by Cambridge University Press

that the proportion of pervasively and situationally hyperactive children in the remaining sample was equal to those estimated for a general population sample. This suggests the sample was representative of the general population and free of any known bias.

To summarise, genetic factors showed widespread influence on problem behaviours in children, as measured by the Conners' scales. These findings refer to normal variation on these behavioural dimensions in the general population. There was no evidence of contrast effects in ratings by teachers, but parent' ratings on several sub-scales showed evidence of such competitive sibling interaction or rater bias effects.

Acknowledgements

We thank Doug Barrett, Emma Canning and Nisha Karia for invaluable help during data collection. We also thank all the parents and teachers who took part in this study. This research was funded by a Wellcome Prize Studentship to the first author.

References

1 Rutter M, Silberg J, O'Connor T, Simonoff E. Genetics and child psychiatry: II empirical research findings. J Child Psychol Psychiatry 1999; 40: 19–55.
2 Edelbrock C, Rende R, Plomin R, Thompson LA. A twin study of competence and problem behaviour in childhood and early adolescence. J Child Psychol Psychiatry 1995; 36: 775–785.
3 Schmitz S, Fulker DW, Mrazek DA. Problem behavior in early and middle childhood: An initial behavior genetic analysis. J Child Psychol Psychiatry 1995; 36: 1443–1458.
4 Dilalla LF, Gottesman II. Heterogeneity of causes for delinquency and criminality: Lifespan perspectives. Dev Psychopathol 1989; 1: 339–349.
5 Eaves LJ, Silberg JL, Meyer JM, Maes HH, Simonoff E, Pickles A, Rutter M, Neale MC, Reynolds CA, Erikson MT, Heath AC, Loeber R, Truett KR, Hewitt J. Genetics and developmental psychopathology: 2. The main effects of genes and environment on behavioral problems in the Virginia Twin Study of Adolescent Behavioral Development. J Child Psychol Psychiatry 1997; 38: 965–980.
6 Neale MC, Cardon LR. Methodology for Genetic Studies of Twins and Families, Kluwer Academic Publishers: Dordrecht. 1992.
7 Nadder TS, Silberg JL, Eaves LJ, Maes HH, Meyer JM. Genetic effects on ADHD symptomatology in 7- to 13-year-old twins: results from a telephone survey. Behav Genet 1998; 28: 83–99.
8 Thapar A, Hervas A, McGuffin P. Childhood hyperactivity scores are highly heritable and show sibling competition effects: Twin study evidence. Behav Genet 1995; 25: 537–544.
9 Saudino KJ, McGuire S, Reiss D, Hetherington EM, Plomin R. Parent ratings of EAS temperaments in twins, full siblings, half siblings, and step siblings. J Pers Soc Psychol 1995; 68: 723–733.
10 Spinath FM, Angleitner A. Contrast effects in Buss and Plomin's EAS questionnaire: A behavioral-genetic study on early developing personality traits assessed through parental ratings. Pers Individ Dif 1998; 25: 947–963.
11 Simonoff E, Pickles A, Hervas A, Silberg J-L, Rutter M, Eaves L. Genetic influences on childhood hyperactivity: Contrast effects imply parental rating bias, not sibling interaction. Psychol Med 1998; 28: 825–837.
12 Goyette CH, Conners CK, Ulrich RF. Normative data on revised Conners' parent and teacher rating scales. J Abnorm Child Psychol 1978; 6: 221–236.
13 Kuntsi J, Oosterlaan J, Stevenson J. Psychological mechanisms in hyperactivity: I Response inhibition deficit, working memory impairment, delay aversion or something else? J Child Psychol Psychiatry, in press.
14 Kuntsi J, Stevenson J. Psychological mechanisms in hyperactivity: II The role of genetic factors. J Child Psychol Psychiatry, in press.
15 Nichols RC, Bilbro WC. The diagnosis of twin zygosity. Acta Genet Med Gemellol (Roma) 1966; 16: 265–275.
16 Cohen DJ, Dibble E, Grawe JM, Pollin W. Reliably separating identical from fraternal twins. Arch Gen Psychiatry 1975; 32: 1371–1375.
17 Plomin R, DeFries JC, McClearn GE, Rutter M. Behavioral Genetics, 3rd edn. W H Freeman: New York, 1997.
18 Neale MC. Mx: Statistical Modelling, 4th edn. Department of Psychiatry: Richmond, VA, 1997.
19 Akaike H. Factor analysis and AIC. Psychometrika, 1987; 52: 317–332.
20 McDonald RP. An index of goodness-of-fit based on noncentrality. J Classif 1989; 6: 97–103.
21 Gjone H, Stevenson J, Sundet JM. Genetic influence on parent-reported attention-related problems in a Norwegian general population twin sample. J Am Acad Child Adolesc Psychiatry 1996; 35: 588–596.
22 Thapar A, McGuffin P. Are anxiety symptoms in childhood heritable? J Child Psychol Psychiatry 1995; 36: 439–447.

https://doi.org/10.1375/twin.3.4.251 Published online by Cambridge University Press