
Twins and Non-twin Siblings: Different Estimates of Shared Environmental Influence in Early Childhood

Gesina Koeppen-Schomerus¹, Frank M. Spinath^{1,2}, and Robert Plomin¹

¹*Social, Genetic and Developmental Psychiatry Research Centre, Institute of Psychiatry, King's College London, UK*

²*Fakultät für Psychologie und Sportwissenschaft, Abteilung für Psychologie, Universität Bielefeld, Bielefeld, Germany*

Twin studies typically indicate shared environmental influence for cognitive abilities, especially in early childhood. However, across studies, DZ twin correlations tend to be greater than non-twin sibling correlations, suggesting that twin estimates of shared environment are to some extent specific to twins. We tested this hypothesis in a sample of more than 1800 MZ and 1800 same-sex DZ pairs from the Twins Early Development Study (TEDS), a study of twins born in England and Wales in 1994 and 1995. For this analysis, we obtained comparable data from more than 130 same-sex younger siblings of the twins. Twins and their younger siblings were assessed for language, cognitive abilities and behavior problems by their parents at 2 and 3 years of age. For language and cognitive measures at both 2 and 3 years, but not for behavior problems, estimates of shared environment were more than twice as large for twins as compared to non-twin siblings. We conclude that about half of twin study estimates of shared environment for cognitive abilities in early childhood are specific to twins. Although many possibilities exist for explaining the special shared environment effect for twins, we suggest that cognitive-relevant experiences that are not shared by siblings are shared by twins because they are exactly the same age.

One of the most important findings from twin and adoption studies concerns nurture rather than nature. Although theories of socialization assume that environments are doled out on a family-by-family basis, behavioral genetic research shows that, after controlling for genetic resemblance, growing up in the same family does not make children similar in personality or psychopathology (Harris, 1998; Plomin & Daniels, 1987). That is, siblings are similar but the reason is nature rather than nurture. Cognitive abilities appear to be an exception to the rule that environmental influences are non-shared rather than shared. A direct estimate of shared environmental influence comes from correlations for adoptive siblings, genetically unrelated children brought together by adoption. The adoptive sibling correlation for full scale IQ for more than 700 pairs of adoptive siblings is .32 (Bouchard & McGue, 1981), suggesting that about one-third of the total variance of IQ scores can be explained by shared environmental influence. Although twin studies are less powerful than the adoptive sibling design for assessing shared environmental influence (Neale & Cardon, 1992), they yield a similar estimate of shared environment for IQ. The average IQ correlation across all studies reviewed by Bouchard et al.

(1981) for more than 10,000 same-sex pairs of twins is .86 for MZ twins and .60 for DZ twins, again suggesting that about one-third of the variance can be explained by shared environmental influence.

Thus, adoption and twin studies appear to tell a consistent story: for cognitive abilities — at least for IQ which has been studied much more than specific cognitive abilities — shared environment accounts for about one-third of the total variance. However, the plot thickens when development is considered. Most of the adoptive siblings in these studies were young with an average age of 10 years. In contrast, four studies of post-adolescent adoptive siblings find an average IQ correlation of $-.01$, suggesting that after adolescence shared environmental influence for IQ is negligible (reviewed in Plomin et al., 1994).

Twin studies, on the other hand, do not indicate a decline in shared environmental influence on IQ from childhood to adolescence. For example, for 95% of twin research on IQ, subjects are from 6 to 20 years of age. Dividing the samples into three age groups (childhood: 6–12 years, early adolescence: 12–16 years, late adolescence: 16–20 years), McGue et al. (1993) report shared environmental estimates of about .30 for all three age intervals. The results further suggest that shared environment declines later in adulthood however, sample sizes are small.

After adolescence, why do twin studies suggest shared environmental influence but studies of adoptive siblings do not? The most likely hypothesis is that twins share IQ-relevant environments to a greater extent than do non-twin siblings. Support for this hypothesis comes from a review of the IQ literature for more than 25,000 pairs of non-twin siblings. In contrast to the average IQ correlation of .60 for DZ twins, the average IQ correlation is .47 for non-twin siblings (Bouchard et al., 1981). Because DZ twins and non-twin siblings are just as similar genetically, the greater IQ resemblance for DZ twins is attributed to greater shared environmental influences. Although no model-fitting analysis has focused specifically on comparing correlations

Address for correspondence: Gesina Koeppen-Schomerus, SGDP Research Centre, Institute of Psychiatry, PO80 London SE5 8AF, United Kingdom. Email: G.Koeppen-Schomerus@iop.kcl.ac.uk

of twins and non-twin siblings, a model-fitting analysis of all family, adoption, and twin studies of IQ considered the possibility of different shared environmental effects for twins and non-twin siblings (Chipuer et al., 1990). This analysis yielded shared environment estimates of about 40% for twins and about 25% for non-twin siblings. The greater shared environment estimate for twins as compared to non-twin siblings was largely due to the difference in correlations between DZ twins and non-twin siblings. However, the shared environment estimate was higher than discussed above (McGue et al., 1993) for both twins (40% versus 30%) and for non-twin siblings (25% versus 0%) because data for adoptive siblings consisted solely of young children in this analysis. A model-fitting analysis of the same data, which attributed excess DZ similarity to prenatal effects, reported an estimate of "prenatal shared environment" of about 20% (Devlin et al., 1997).

Although comparisons between twin and non-twin sibling results across different studies support the hypothesis that during childhood twins share environments relevant to IQ to a greater extent than do non-twin siblings, more definitive data come from examining twin and non-twin correlations within the same study. The Louisville Twin Study (LTS; Wilson, 1983) assessed 74 age- and test-matched siblings of 129 twin pairs at 3 years. The twins and their siblings were studied longitudinally from infancy to adolescence. In early childhood (age 2 to 4), IQ correlations between the siblings and their twins (which we shall call twin-sibling correlations) were only half the magnitude of DZ correlations, thus supporting the comparison of twin and non-twin results across different studies. After age 4, the differences between twin-sibling correlations and DZ twin correlations were notably reduced. Another clue in the puzzle of twin study estimates of shared environment is that the LTS suggested substantial shared environment in infancy and early childhood, accounting for over 60% of the IQ variance. In middle childhood (age 5 to 9), shared environmental estimates averaged .41 and decreased to .20 in adolescence which is in line with the results described above.

The purpose of the present study was to confirm the LTS results for IQ using a large sample and model-fitting techniques that explicitly compared twin and twin-sibling results. In addition, we extended the twin and twin-sibling comparison to consider key developmental variables in addition to IQ: language development, nonverbal-cognitive development and behavior problems.

Method

Sample

The Twins Early Development Study (TEDS) is a national study of twins born in England and Wales from 1994 to 1996 (Trouton et al., 2002). All families who indicated their interest in participating in the study (71%, $N = 11,352$) were sent a questionnaire booklet requesting background information about the family when the twins were a year and a half old. The initial background booklet which asks parents about family details was completed and returned by 9206 families (81%). Shortly before the twins' second, third and fourth birthdays postal questionnaires and assessment booklets were sent to the parents to assess their twins' cognitive

abilities, language abilities and behavior problems. Informed consent was obtained at each stage. From this sample, a completed 2-year booklet was returned by 6137 (67%) families, a completed 3-year booklet was returned by 5909 families (64%), and 4,986 families (54%) completed the booklets sent out at age 4. For more than 97% of TEDS families, the questionnaire and assessment booklets were completed by the mothers.

Despite the attrition it has been demonstrated that the present TEDS sample is reasonably representative of the UK population of parents of young children in terms of education, parental ethnicity, and employment status as compared to 1994 census data from the Office of National Statistics (Price et al., 2001).

Physical similarity ratings by parents were used to determine the zygosity of the twins, a method which assigns zygosity with 95% accuracy as assessed using DNA markers (Price et al., 2000). Regarding the potential misclassification of MZ twins as DZ, which would artificially inflate DZ twin correlations, MZ pairs were no less likely than DZ pairs to be misclassified (Price et al., 2000).

Families were excluded if zygosity was uncertain (184 families), or if one or other twin had severe medical or genetic problems (179 families) or severe perinatal problems (188 families). The exclusion criteria also involved incomplete booklets or booklets returned later than 6 months after the twins' birthday, and if English was not the first language spoken in the home.

In addition to the twin data, TEDS also asked parents whether the twins have any younger brothers and sisters, and whether they would like them to participate in the study as well. To date more than 1000 younger siblings are participating in TEDS. The same questionnaire booklets that were sent to the twins when they were 2, 3 and 4 years old were also sent to their younger siblings when they reach these ages allowing an age-matched comparison between the younger siblings and the twins. The total number of younger siblings with available data at ages 2 and 3 was 427. Of these, 303 (71%) completed the booklets at 2 years and 283 (66.3%) returned the assessments at age 3. Data at 4 years were not used due to an insufficient number of younger siblings assessed at this age at the time of analysis. Due to the restricted sample size within the younger sibling group, we focused on same-sex twin and same sex twin-sibling pairs only.

Twins and twin-sibling pairs were limited to those with complete data at 2 and 3 years. At age 2, the target sample used in the present study consisted of 3891 families from the 1994 and 1995 cohort, including 1865 MZ twin pairs (MZ), 1896 same-sex DZ twin pairs (DZss), and 130 same sex twin-non-twin sibling pairs (TS). At age 3, the target sample included 3743 families, made up of 1805 MZ, 1801 DZss and 137 TS pairs.

In order to create independent groups, the MZ and DZ groups were selected to include twins who did not have a younger sibling. For the twin-sibling pairs (TS) one member from a twin pair with a younger sibling of the same sex was selected at random and paired with his/her younger brother or sister. Younger siblings who did not

share the same parents as their twin siblings were excluded from the analyses.

Measures

Assessment of Verbal Ability

Verbal performance was assessed using composite scores derived from age-appropriate vocabulary and grammar scales from the MacArthur Communicative Development Inventory: UK Short Form. This measure is an abbreviated and anglicised adaptation of the MacArthur Communicative Development Inventory (MCDI), a widely-used measure of early language development (Dale et al., 1998). The MCDI has been shown to have excellent internal consistency and test-retest reliability, as well as concurrent validity with tester-administered measures (Fenson et al., 2000).

Vocabulary production was assessed by means of a 100 item-checklist asking parents to report on their children's production of root words (e.g., dog, game, gentle). A composite score was calculated by summing the number of words checked. The vocabulary measure given at 2 years consisted of a list of 100 words that predicts the 680-word original version of the MCDI with very high accuracy ($r = .98$) (Fenson et al., 2000). The 3-year measure was a similar 100 item-checklist, taken from an upward extension of the MCDI.

Grammar was assessed in the MCDI by asking parents about their children's sentence complexity. The grammar scales at 2 and 3 years each comprised 13 questions. The first question asked whether the child is able to combine words, the remaining 12 items each present two sentences carrying the same meaning, with the first representing a developmentally simpler form (e.g., "I want truck" vs. "I want truck like Billy has"). Parents indicated the sentence in each pair that resembled most what their child was able to say. Scoring details are reported by (Price et al., 2001).

The MCDI vocabulary and grammar scores at 2 and 3 years were combined in a composite measure at each year which we refer to as the MCDI verbal score. Details of the construction of the composite are described elsewhere (Spinath et al., 2003).

Assessment of Nonverbal Cognitive Ability

Nonverbal cognitive ability was assessed using age-appropriate versions of the Parent Report of Children's Abilities (PARCA; Oliver et al., 2001; Saudino et al., 1998). The PARCA is an hour-long test developed from various standard measures of cognitive ability (e.g., McCarthy Scales of Children's Abilities; McCarthy, 1972; Bayley Scales of Infant Development; Bayley, 1993), as well as novel items designed specifically for the PARCA. It is made up of several subtests within two sections. In the first section, parents administer a battery of standard cognitive tasks to their children. For the 2-year version of the PARCA, for example, the parent-administered section consists of block building, drawing designs, design copying, matching, and paper-folding tasks. In the second section of the PARCA, parents report on the ability of their children to perform specific cognitive tasks (e.g., "Does your child recognise himself/herself when looking in the mirror?"). Age-appropriate versions of the 2-year PARCA were used at age 3.

Details of the test material are provided by Price et al. (2001). Validation studies at ages 2 and 3 have demonstrated good internal consistencies for both the parent-administered and parent-reported, as well as reasonable validity as assessed in the home using traditional tester-administered measures (Oliver et al., 2001; Saudino et al., 1998).

In the present analyses, we used a composite measure at 2 and at 3 years that is derived from the parental-administered and parent-report sections of the PARCA. Details of the construction of the composite are described elsewhere (Spinath et al., 2003). We refer to this composite as the PARCA nonverbal score.

Assessment of General Cognitive Ability ("g")

Principal component analyses were applied to age- and sex-corrected data from the PARCA and the MCDI. At each age, the first unrotated component accounted for more than 50% of the total variance and was consequently used as an indicator of "g". Further details about the derivation of this component of general cognitive ability are described elsewhere (Spinath et al., 2003).

Assessment of Behavior Problems

Behavior problems were assessed at 2 and 3 years using the Revised Rutter Parent Scale for Preschool Children (RRPSPC; Hogg et al., 1997). The RRPSPC is based on the Preschool Behavior Questionnaire (Behar & Stringfield, 1974), a downward extension to the pre-school range of the original Rutter Parent Scale (Rutter et al., 1970). The test-retest reliability of the original Rutter Parent Scale total behavior problem score was 0.87 (Rutter et al., 1970). The behavior problems scales include 44 and 52 items at 2 and 3 years respectively. In this analysis, we report results for the total problems score at 2 and 3 years, which we refer to as the RRPSPC score.

The analyses reported here are based on age and sex corrected scores for all measures.

Analyses

The essence of the design is very simple. If twins share environments to a greater extent than non-twin siblings, we would expect correlations for DZ twin pairs to exceed correlations for twin-sibling pairs. In order to test the quantitative effects of twin-specific environments, structural equation modelling was applied to the variance/covariance matrices for the four scales at 2 and 3 years. We used Mx (Neale et al., 1999) to test the two models shown in Figure 1.

In Model 1 the parameter estimates A, C and E have been equated for twins and twin-sibling pairs. For twin-sibling pairs, the common environmental correlation (rc) is either freed or fixed to 1.0. A likelihood ratio χ^2 test (LRT) assesses the significance of the χ^2 difference when rc is freed or fixed to 1.0 (with $\Delta df = 1$). This allows a comparison of differences in common environmental influence between DZ twins and twin-sibling pairs.

Model 2 estimates the contributions of additive genes (A), shared environment (S) and non-shared environment (E) for both twin-twin and twin-sibling groups. In contrast to model 1, shared environment (S) refers to environmental factors that are shared by children, regardless of whether

they are twins or not. In addition, Model 2 includes a direct estimate of the influence of twin-specific shared environment (T). The model assumes that T contributes to the covariance of twin-twin pairs but not of twin-sibling pairs. In twin-sibling pairs, T is subsumed by non-shared environmental variance. Also, the common environment term (C) of Model 1 corresponds to the sum of shared environment (S) plus twin specific shared environment (T).

There are of course other ways in which these data could be analysed. For example, raw data approaches could be used instead of variance/covariance matrices. Also, with a larger number of siblings the sample could have been divided into four groups — MZ twins, DZ twins, MZ twins plus sibling, DZ twins plus sibling. At present, data collection in TEDS for the twins' younger siblings (at ages 2, 3 and 4 years) is still ongoing and we anticipate applying other models in future analyses. However, given that the basic results we report can be seen simply in the twin and twin-sibling correlations, we are confident that similar results will emerge when more complicated models are applied.

Results

The sample sizes and basic descriptive statistics for monozygotic (MZ), dizygotic same-sex (DZss) twin pairs and same-sex twin sibling pairs (TS) are shown in Table 1 for general cognitive ability ("g"), verbal ability (MCDI), nonverbal cognitive ability (PARCA), and behavior problems (RRPSPC) at ages 2 and 3. We also conducted one-way analyses of variance (ANOVA) to test for possible significant differences of means and variances between MZ, DZ and TS groups and have summarised the results in Table 1. We used Levene's test to assess homogeneity of variance between the three groups. Where significant differences between variances were observed, we applied Dunnett's C test for post hoc comparisons of mean differences between groups; if no significant differences between variances were found, Tukey's HSD test was used.

Post hoc comparisons showed significant mean differences between MZ and DZ groups for most outcome measures, whereas for MZ-TS and DZ-TS comparisons, significant mean differences were less frequently observed and *p*-values were generally found to be lower. Variance differences were significant for "g" at both ages (*p* < .05) and, at age 3 only, for the MCDI and the PARCA (*p* < .05 and *p* < .001, respectively). For all cognitive measures at age 3 years, a pattern could be observed where the variances of twin-twin pairs exceeded those of twin-sibling pairs (with $\text{var}_{\text{MZ}} > \text{var}_{\text{DZ}} > \text{var}_{\text{TS}}$). Rather than further exploring mean differences between twins and singletons (e.g., Posthuma et al., 2000), the present analyses focus on the comparison of twin and non-twin sibling correlations to estimate shared and non-shared environmental components of variance.

Intraclass correlations were obtained for twin-twin and twin-sibling pairs on all four measures at ages 2 and 3. As shown in Table 2, MZ twin correlations were consistently and significantly higher than DZ correlations for all four measures, suggesting genetic influence. The critical comparison in the present context is the comparison between TS and DZ correlations. For the three cognitive measures,

TS correlations were considerably lower than DZ correlations, suggesting a twin-specific shared environmental effect. For behavior problems (RRPSPC), on the other hand, the correlations were similar in magnitude for DZ (at 2 years: *r* = .48; at 3 years: *r* = .47) and TS pairs (at 2 years: *r* = .43, at 3 years: *r* = .42).

Because non-twin siblings differ in age on average by about 2 years, they can be assessed at the same age as in TEDS but they cannot be tested at the same measurement occasion as can twins. This is of course a possible reason why twin correlations exceed non-twin-sibling correlations. An interesting point of comparison is the correlation from 2 to 3 years because these stability correlations create a ceiling for the twin-sibling correlations. For all individuals, the correlations from 2 to 3 years are .68 for "g", .61 for MCDI, .57 for PARCA, and .62 for RRPSPC. These stability correlations from 2 to 3 years are lower than the twin correlations, which suggests that testing twins simultaneously could contribute to twin-specific shared environments.

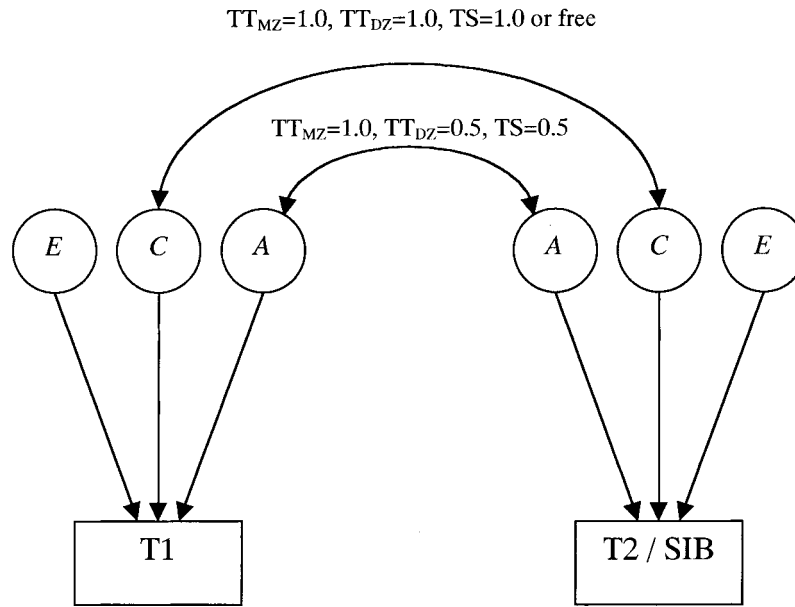
Model-fitting results and parameter estimates are summarised in Tables 3 a) and b). For all cognitive measures ("g", MCDI and PARCA), model 1 fitted the data significantly better (*p* < .001) when the common environmental correlation (*r_c*) for TS pairs was free. The results were very highly similar at both ages. The common environmental correlation (*r_c*) for cognitive scales ranged from .50 to .68 for TS pairs as compared to 1.0 for twin-twin pairs.

We used model 2 to explore further the differences in common environment between twin-twin and twin-sibling pairs. Applying model 2 to the general cognitive ability ("g") data, we found additive genetic effects to be moderate (22–30%) and shared environmental influences (S) around 40%. In addition, we found a significant effect of twin specific environment (T) which accounted for 20–31% of the variance. The overall effect of common environment (i.e., $C = S + T$) explained between 60–70% of the variance.

For measures of verbal and nonverbal cognitive ability, a similar pattern of results emerged. Heritability estimates were again modest (between 20–25%) across both scales and ages. Twin-specific shared environment effects (T) accounted for a significant proportion of the variance at ages 2 and 3 for the MCDI (33% and 26%) and the PARCA (30% and 28%).

In contrast to the consistent findings showing twin-specific shared environment effects for our cognitive measures, we found no decrease in the fit of model 1 for data from the behavior problems scale (RRPSPC) when the environmental correlation (*r_c*) between TS pairs was fixed to 1.0. Additive genetic factors explained 57% and 55% of the variance at ages 2 and 3. Common environmental effects accounted for almost 20% of the variance at both ages and non-shared environmental effects explained 24% and 26% of the variance at 2 and 3 years, respectively. These findings for the RRPSPC indicate an absence of twin-specific environmental effects for behavior problems. For all cognitive measures, on the other hand, we consistently demonstrated significant twin-specific shared environmental effects at both ages ranging from 20–33%.

Model 1:



Model 2:

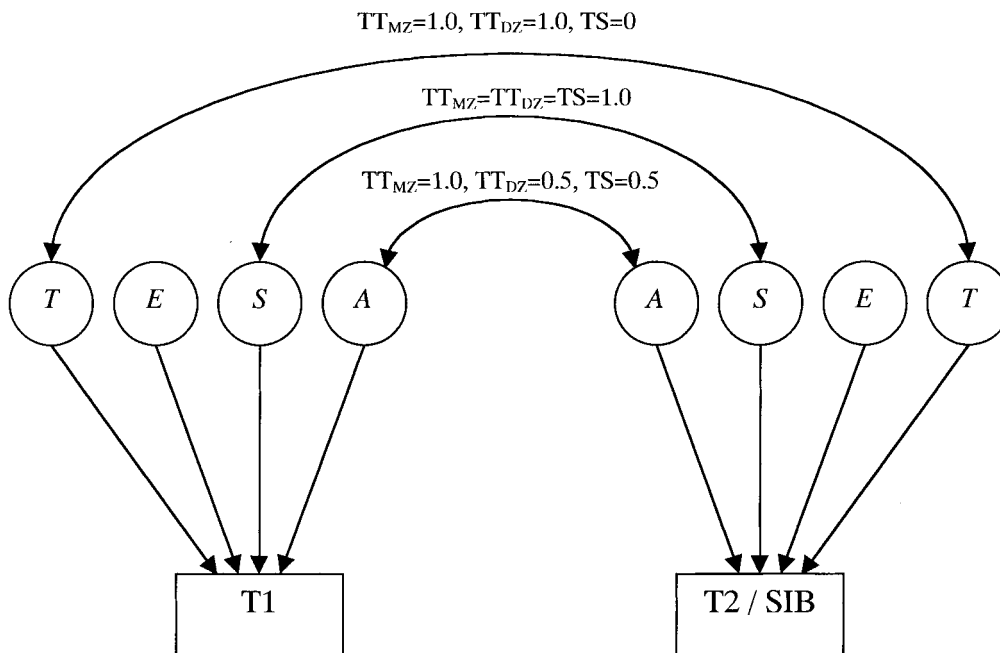


Figure 1

Modelling of specific twin environment in a mixed twin and sibling sample.

Model 1 decomposes the variance into additive genetic (A), common environmental (C) and non-shared environmental components (E) for both twin-twin (MZ and DZ) and twin-sibling (TS) groups. The shared environmental correlation (r_c) for twins is by definition 1.0. For TS pairs r_c is either fixed to 1.0 or freed and the difference in fit for the model is observed.

Model 2 decomposes the variance into additive genetic (A), shared environment (S) and non-shared environment (E) for both groups. Instead of estimating r_c , the model introduces an additional environmental parameter, T (relating to twin-twin specific shared environmental variance) which contributes to the covariance of twin-twin pairs but not to the covariance of twin-sibling pairs. In twin-sibling pairs, T reflects non-shared environmental variance. In addition, the common environment term (C) of model 1 is here the sum of shared environment (S) which is shared by siblings as well as twins and twin specific shared environment (T).

Table 1

Sample Sizes, Basic Descriptive Statistics and Analysis of Variance (ANOVA) Results for Monozygotic (MZ), Dizygotic Same Sex (DZss) Twin Pairs and Twin-sibling Pairs (TS) for General Cognitive Ability (“g”), Verbal Ability (MCDI), Nonverbal Cognitive Ability (PARCA), and Behavior Problems (RRPSPC) at Ages 2 and 3 Years

Measure	Age	MZ		DZ		TS		ANOVA						
		N (pairs)	mean	SD	N (pairs)	mean	SD	N (pairs)	mean	SD	MZ-DZ	MZ-TS	DZ-TS	Levene's Test
“g”	2	1297	-0.07	1.02	1302	0.07	.98	88	-0.03	.99	< .001	<i>ns</i>	<i>ns</i>	< .05
	3	1277	-0.04	1.02	1303	0.05	.98	51	-0.16	.89	< .05	<i>ns</i>	<i>ns</i>	< .05
MCDI	2	1685	-0.08	1.00	1695	0.07	.99	111	0.10	1.06	< .001	< .05	<i>ns</i>	<i>ns</i>
	3	1493	-0.05	1.03	1522	0.04	.98	100	0.10	.94	< .001	<i>ns</i>	<i>ns</i>	< .001
PARCA	2	1735	-0.04	1.01	1740	0.03	.99	111	0.19	.91	< .05	< .05	< .05	<i>ns</i>
	3	1693	-0.01	1.01	1705	0.00	.99	122	0.18	.90	<i>ns</i>	< .05	< .05	< .05
RRPSPC	2	1717	-0.04	.99	1731	0.07	1.01	116	-0.23	.99	< .001	< .05	< .001	<i>ns</i>
	3	1702	-0.03	.99	1716	0.04	1.01	126	-0.14	.99	< .05	<i>ns</i>	< .05	<i>ns</i>

Note: For ANOVAs where Levene's test for homogeneity of variance was significant, post-hoc comparisons of mean differences between groups are based on Dunnett's C test; Tukey's HSD test was used when there was no significant difference between variances. *ns* = not significant.

Table 2

Intraclass Correlations and 95% Confidence Intervals (CI) for Monozygotic (MZ), Dizygotic Same Sex (DZss) Twins and Twin-sibling (TS) Pairs for General Cognitive Ability (“g”), Verbal Ability (MCDI), Nonverbal Cognitive Ability (PARCA), and Behavior Problems (RRPSPC) at Ages 2 and 3 Years

	2 years				3 years			
	“g”	MCDI	PARCA	RRPSPC	“g”	MCDI	PARCA	RRPSPC
r_{MZ}	0.93	0.96	0.83	0.75	0.92	0.91	0.85	0.73
CI	.92-.94	.95-.96	.82-.85	.73-.77	.91-.93	.91-.92	.83-.86	.71-.75
r_{DZss}	0.81	0.85	0.72	0.48	0.76	0.79	0.72	0.47
CI	.79-.83	.83-.86	.69-.74	.44-.52	.73-.78	.77-.81	.69-.74	.43-.50
r_{TS}	0.47	0.56	0.35	0.43	0.49	0.50	0.35	0.42
CI	.28-.62	.41-.67	.18-.50	.22-.59	.26-.67	.34-.63	.18-.49	.25-.56

Table 3a

Model Fitting Results: Fit Statistics and Parameter Estimates for General Cognitive Ability (“g”), Verbal Ability (MCDI), Nonverbal Cognitive Ability (PARCA) and Behavior Problems (RRPSPC) Assessed at Ages 2 and 3 — Model Fitting Results at Age 2 Years

	Models 1 & 2		Model 1			Model 2			
	Model fit		$r_c = \text{free}$	$r_c = 1.0$		Parameter Estimates (95% CI)			
	χ^2	p	r_c	LRT	p	a^2	s^2	t^2	e^2
“g”	2.93	.71	.57 (.33-.74)	53.30	< .001	.22 (.19-.25)	.40 (.23-.52)	.31 (.19-.47)	.07 (.06-.08)
MCDI	3.41	.64	.56 (.38-.67)	114.26	< .001	.20 (.18-.23)	.42 (.29-.52)	.33 (.23-.47)	.04 (.04-.05)
PARCA	7.9	.16	.50 (.19-.73)	31.52	< .001	.22 (.17-.26)	.31 (.12-.45)	.30 (.17-.49)	.17 (.16-.19)
RRPSPC	4.41	.49	1.0 (.77-1.0)	0.00	<i>ns</i>	.57 (.50-.64)	.19 (.06-.26)	.00 (.00-.13)	.24 (.22-.26)

Discussion

The findings of the present study suggest the existence of twin-specific effects for estimates of shared environment for measures of cognitive abilities. Our results show that for cognitive measures (general cognitive ability “g”, verbal ability and nonverbal cognitive ability) in early childhood, a considerable proportion of variance can be attributed to shared environmental effects that are twin-specific (20–33%), in addition to those that are not specific to twins (31–43%).

In line with findings reported by the Louisville Twin Study (LTS; Wilson, 1983), we found differences in common environment for “g” between twins and their non-twin siblings. Figure 2 shows that the present TEDS results at 2 and 3 years are highly similar to those from the LTS. This similarity is particularly noteworthy, because the two studies assessed cognitive ability in childhood using different measurement methods — parent-administered measures in TEDS and tester-administered measures in

Table 3b

Model Fitting Results: Fit Statistics and Parameter Estimates for General Cognitive Ability (“g”), Verbal Ability (MCDI), Nonverbal Cognitive Ability (PARCA) and Behavior Problems (RRPSPC) Assessed at Ages 2 and 3 — Model Fitting Results at aAge 3 Years

	Models 1&2		Model 1			Model 2			
	Model fit		$r_c = \text{free}$	$r_c = 1.0$		Parameter Estimates (95% CI)			
	c^2	p	r_c	LRT	p	a^2	s^2	t^2	e^2
“g”	13.82	.02	.68 (.30–.91)	9.90	< .01	.30 (.26–.34)	.42 (.18–.56)	.20 (.06–.43)	.08 (.08–.09)
MCDI	9.22	.10	.62 (.39–.78)	39.91	< .001	.22 (.19–.26)	.43 (.27–.54)	.26 (.15–.42)	.09 (.08–.10)
PARCA	6.53	.26	.52 (.22–.74)	29.64	< .001	.25 (.21–.30)	.31 (.13–.44)	.28 (.15–.46)	.16 (.14–.17)
RRPSPC	2.93	.71	.98 (.19–1.0)	0.001	<i>ns</i>	.55 (.48–.63)	.18 (.03–.25)	.00 (.00–.15)	.26 (.24–.28)

Note: r_c = common environmental correlation for TS pairs; a^2 = additive genes; c^2 = common environment; s^2 = shared environment; t^2 = twin-specific shared environment; e^2 = non-shared environment; The common environment term (C) in Model 1 corresponds to the sum of shared environment (S) which is shared by siblings as well as twins plus twin specific shared environment (T).

In Model 1, the common environmental correlation (r_c) for TS pairs is either freed or fixed to 1.0. LRT is the Chi-square difference between those two variants of Model 1 (with $Df = 1$). In Model 2, t^2 is estimated directly. The fit indices are identical for Model 2 and the variant of Model 1 with r_c freed ($Df = 5$).

LTS. In both studies, correlations for twin-sibling pairs were substantially lower than correlations for DZ twins.

For behavior problems, on the other hand, we found similar correlations for DZ twins and twin-siblings. In other words, there was no indication of twin-specific shared environmental effects on behavior problems.

For “g”, we estimated shared environmental influences to be about 65% for twins (i.e., S + T) and about 40% for twin-siblings. Our results suggest that the common environment estimates derived from our twin sample alone include an additional 25% (i.e., the difference between C and S which is across ages 65%–40%) due to a special shared environment effect for twins (T). Furthermore, as suggested by IQ data for adoptive siblings (Chipuer et al., 1990), the shared environmental estimate of 40% for our twin-siblings should decline after childhood. We predict that the 65% twin estimate for common environment will decline to 25% after childhood. The remaining 25% exclusively represents a special shared environment effect for

twins, as seen in twin studies of IQ in adolescence and adulthood (McGue et al., 1993).

An important next step involves a better understanding of the processes involved in the special shared environmental twin effect on cognition. One hypothesis to explain the greater DZ twin resemblance than twin-sibling resemblance involves prenatal effects because twins share the womb at the same time. Although prenatal effects appear to contribute to the resemblance of DZ twins throughout early childhood, these effects tend to diminish by middle childhood (Wilson, 1979). On the other hand, it has been argued that prenatal effects contribute more to post-natal differences between twins rather than to their increased similarity (Resnick et al., 1993; Stauffer et al., 1998). If prenatal effects diminish in childhood and if the special shared environmental twin effect persists from childhood into adulthood, the prenatal hypothesis seems less likely to explain the special shared environmental twin effect.

Assortative mating might seem to be a better hypothesis for explaining the special shared environmental twin effect for two reasons: it is substantial for cognitive abilities but not for personality or psychopathology and its effect could be constant throughout development. Spouse correlations for IQ are about .40 (Jensen, 1978) in contrast to about .15 for personality (Vandenberg, 1972). To the extent that assortative mating operates genetically, it will increase genetic resemblance for DZ twins but not for MZ twins because MZ twins are genetically identical. In this way, assortative mating inflates twin study estimates of shared environment. However, assortative mating cannot explain why twin-sibling correlations are lower than DZ twin correlations because it should have the same effect on DZ twins and twin-siblings. Thus, although model-fitting twin analyses that take assortative mating into account yield reduced estimates of shared environment, assortative mating cannot explain the special effect of shared environment for twins.

Age differences within pairs of non-twin siblings are another possible source for lower similarity. The average IQ correlation of 0.47 for siblings (Bouchard et al., 1981) involves members of sibling pairs who were assessed at different ages, whereas for the average DZ correlation of 0.60,

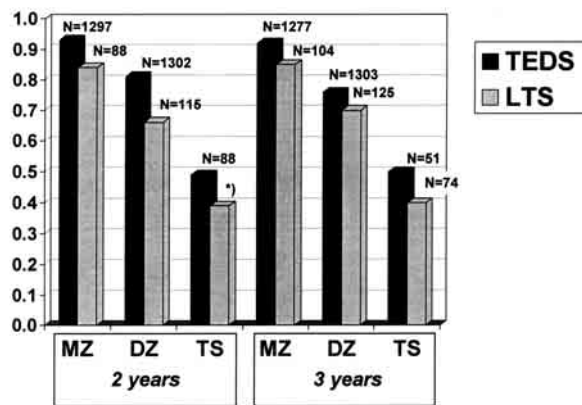


Figure 2

Comparison of Intraclass correlations for MZ, DZ and TS pairs at ages 2 and 3 for general cognitive ability (“g”) between the Louisville Twin Study (LTS) and the Twins Early Development Study (TEDS).

Note: LTS correlations are based on data published by Wilson (1983)

*) details not available

both twins were tested at exactly the same age. The greater age gap for non-twin siblings could be responsible for the lower correlations for non-twin siblings than DZ twins. During mid-childhood or adolescence, this is not a strong hypothesis to explain the special twin effect because IQ scores are age-standardized and residual age effects on twin resemblance are weak. In early childhood though, age differences within pairs of non-twin siblings could be a better hypothesis because the age gap between young children and their older siblings represents a large proportion of their entire life. However, in the longitudinal designs of LTS and TEDS, the younger siblings of the twins were tested at the same age as the twins, which eliminates age differences as an explanation of the lower twin-sibling correlations.

The remaining set of hypotheses involves postnatal experiences. The first hypothesis is that there are special twin-specific shared environmental processes that make twins more similar than non-twin siblings. For example, parents and teachers might treat twins more similarly. One problem with this hypothesis is that parent and teacher effects seem unlikely to explain persistence of the special C twin effect from childhood to adulthood when parents and teachers presumably have less effect. A related postnatal experiences hypothesis that does not have this problem focuses on the twin relationship: one constant in twins' lives is their twin partner. Perhaps twins have special environmental effects on each other that do not affect non-twin siblings. Both of these postnatal hypotheses assume that there are processes that uniquely and substantially affect twins, which implies that variances for twins would be expected to exceed variance for non-twin siblings. Furthermore, the greater variance for twins than non-twin siblings should be seen for cognitive measures (which show the special twin effect) but not for behavior problems (which do not show the special twin effect). Comparisons of variances for twins and non-twin siblings provide some support for this hypothesis at 3 years (Table 1), although caution is warranted because our current sample of non-twin siblings is small relative to the size of the twin samples. If correct, these postnatal hypotheses imply that even if systematic sources of shared environment specific to twins could be identified, they would not be applicable to non-twin siblings.

The default hypothesis that we favour is that the special twin effect involves postnatal experiences that are normally not shared by siblings but are experienced as shared environmental influences by twins because twins are exactly the same age. This hypothesis differs from the previous postnatal hypothesis in that it posits a shifting of non-shared variance to shared variance for twin pairs because they are exactly the same age. In contrast, the previous postnatal hypotheses assume new sources of variance that would increase the variance for members of twin pairs. In other words, the present hypothesis does not require that the total variance for DZ twins or twins in TS pairs exceeds variance for non-twin siblings. If correct, this hypothesis would imply if systematic sources of shared environment specific to twins could be identified, they would be applicable as nonshared environment for non-twin siblings because the same processes are involved but are experienced

at the same time by twins. However, it is possible that many chance experiences that contribute to non-shared environment for non-twin siblings are experienced as shared environmental influences by twins. Another variant of this hypothesis is that the twin-specific effect is an artefact of testing the twins at the same time. Supporting this hypothesis is the finding that the stability correlation from 2 to 3 years was lower than the twin correlations. The stability creates a ceiling for twin-sibling correlations because although they were tested at the same age the interval between their tests was about 2 years. Although non-twin siblings cannot be tested at the same time and the same age, it would be useful to assess the effect of testing twins on different days.

The "same-age" hypothesis does not easily explain all of the facts. For example, as twins grow up and are no longer living in the same family or going to the same school, their worlds would be less shared, which would predict a reduction in the special twin effect in adulthood, but that does not appear to be the case. Another way in which this hypothesis as well as the others fall short is that they cannot easily explain why the special twin effect of shared environment occurs for cognitive abilities but not for behavior problems. A final problem is that the two postnatal hypotheses of special shared environmental twin effects would have to predict that adult DZ twins adopted-apart are considerably less similar than DZ twins reared together. However, results from two studies of adult twins reared apart (Bouchard et al., 1990; Pedersen et al., 1992) yield correlations that are slightly *greater* for DZ twins adopted-apart (r_{DZA} between .32 and .34) than for DZ twins reared together ($r_{DZT} = .22$), although the sample sizes are not nearly large enough to provide adequate power to detect differences.

In summary, our results provide strong support for the hypothesis that, in early childhood, twin-twin estimates of shared environment are about twice as great as estimates based on twin-sibling for cognitive abilities but not for behavior problems. However, the explanation of this special twin effect must await further research.

Acknowledgments

We thank the parents of the twins in the Twins Early Development Study (TEDS) for making the study possible. TEDS is supported by a program grant from the UK Medical Research Council (MRC).

References

- Bailey, N. (1993). *Bailey Scales of Infant Development*. New York: The Psychological Corporation.
- Behar, L., & Stringfield, S. (1974). A behavior rating scale for the preschool child. *Developmental Psychology*, *33*, 3–66.
- Bouchard, T. J., Jr., Lykken, D. T., McGue, M., Segal, N. L., & Tellegen, A. (1990). Sources of human psychological differences: The Minnesota Study of Twins Reared Apart. *Science*, *250*, 223–228.
- Bouchard, T. J., Jr. & McGue, M. (1981). Familial studies of intelligence: A review. *Science*, *212*, 1055–1059.

- Chipuer, H. M., Rovine, M. J., & Plomin, R. (1990). LISREL modelling: Genetic and environmental influences on IQ revisited. *Intelligence, 14*, 11–29.
- Dale, P. S., Reznick, J. S., & Thal, D. J. (1998). *A parent report measure of language development for three-year-olds*. International Conference on Infant Studies, Atlanta, Georgia.
- Devlin, B., Daniels, M., & Roeder, K. (1997). The heritability of IQ. *Nature, 388*, 468–471.
- Fenson, L., Pethick, S., Renda, C., Cox, J. L., Dale, P. S., & Reznick, J. S. (2000). Short-form version of the MacArthur Communicative Development Inventories. *Applied Psycholinguistics, 21*, 95–116.
- Harris, J. R. (1998). *The nurture assumption: Why children turn out the way they do*. New York: The Free Press.
- Hogg, C., Rutter, M., & Richman, N. (1997). Emotional and behavioral problems in children. In I. Inslare (Ed.), *Child psychology portfolio*. Windsor: NFER-Nelson.
- Jensen, A. R. (1978). Genetic and behavioral effects of nonrandom mating. In R. T. Osbourne, C. E. Noble, & N. Weyl (Eds.), *Human variation: The biopsychology of age, race, and sex* (pp. 51–105). New York: Academic Press.
- McCarthy, D. (1972). *McCarthy Scales of Children's Abilities*. New York: The Psychological Corporation.
- McGue, M., Bouchard, T. J., Jr., 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.
- Neale, M. C., Boker, S. M., Xie, G., & Maes, H. (1999). *Mx: Statistical modeling* (5th ed.). VCU Box 900126, Richmond, VA 23298: Department of Psychiatry.
- Neale, M. C. & Cardon, L. R. (1992). *Methodology for genetic studies of twins and families*. Dordrecht: Kluwer Academic Publications.
- Oliver, B., Dale, P. S., Saudino, K., Petrill, S. A., Pike, A., & Plomin, R. (2002). The validity of parent-based assessment of nonverbal cognitive abilities of three-year olds. *Early Child Development & Care, 172*, 337–348.
- Pedersen, N. L., Plomin, R., Nesselroade, J. R., & McClearn, G. E. (1992). A quantitative genetic analysis of cognitive abilities during the second half of the life span. *Psychological Science, 3*, 346–353.
- Plomin, R., Chipuer, H. M., & Neiderhiser, J. M. (1994). Behavioral genetic evidence for the importance of nonshared environment. In E. M. Hetherington, D. Reiss, & R. Plomin (Eds.), *Separate social worlds of siblings: The impact of non-shared environment on development* (pp. 1–31). New Jersey: Lawrence Erlbaum Associates.
- Plomin, R., & Daniels, D. (1987). Why are children in the same family so different from each other? *Behavioral and Brain Sciences, 10*, 1–16.
- Posthuma, D., De Geus, E. J. C., Bleichrodt, N., & Boomsma, D. I. (2000). Twin-singleton differences in intelligence? *Twin Research, 3*, 83–87.
- Price, T. S., Freeman, B., Craig, I. W., Petrill, S. A., Ebersole, L., & Plomin, R. (2000). Infant zygosity can be assigned by parental report questionnaire data. *Twin Research, 3*, 129–133.
- Price, T. S., Petrill, S. A., Dale, P. S., Eley, T. C., & Plomin, R. (2001). *Modularity and genetic correlation in cognitive development: A study of 2- and 3-year-old twins*. Manuscript submitted for publication.
- Resnick, S. M., Gottesman, I. I., & McGue, M. (1993). Sensation seeking in opposite-sex twins: An effect of prenatal hormones? *Behavior Genetics, 23*, 323–329.
- Rutter, M., Tizard, J., & Whitmore, K. (1970). *Education, health and behavior*. London: Longman.
- Saudino, K. J., Dale, P. S., Oliver, B., Petrill, S. A., Richardson, V., Rutter, M., et al. (1998). The validity of parent-based assessment of the cognitive abilities of two-year-olds. *British Journal of Developmental Psychology, 16*, 349–363.
- Spinath, F. M., Ronald, A., Harlaar, N., Price, T. S., & Plomin, R. (2003). Phenotypic “g” early in life: On the etiology of general cognitive ability in a large population sample of twin children aged 2 to 4 years. *Intelligence, 31*, 195–216.
- Stauffer, A., Burns, W. J., Burns, K. A., Melamed, J., & Herman, C. E. (1998). Early developmental progress of preterm twins discordant for birthweight and risk. *Acta Geneticae Medicae et Gemellologiae, 37*, 81–87.
- Trouton, A., Spinath, F. M., & Plomin, R. (2002). Twins Early Development Study (TEDS): A multivariate, longitudinal genetic investigation of language, cognition and behavior problems in childhood. *Twin Research, 5*, 444–448.
- Vandenberg, S. G. (1972). Assortative mating, or who marries whom? *Behavior Genetics, 2*, 127–157.
- Wilson, R. S. (1979). Twin growth: Initial deficit, recovery, and trends in concordance from birth to nine year. *Annals of Human Biology, 6*, 205–220.
- Wilson, R. S. (1983). The Louisville Twin Study: Developmental synchronies in behavior. *Child Development, 54*, 298–316.