

Origins of Individual Differences in Theory of Mind: From Nature to Nurture?

Claire Hughes
University of Cambridge

Sara R. Jaffee
King's College London and University of Pennsylvania

Francesca Happé and Alan Taylor
King's College London

Avshalom Caspi and Terrie E. Moffitt
*King's College London and University of Wisconsin,
Madison*

In this study of the origins of individual differences in theory of mind (ToM), the Environmental Risk (E-Risk) Longitudinal Twin Study sample of 1,116 sixty-month-old twin pairs completed a comprehensive battery of ToM tasks. Individual differences in ToM were striking and strongly associated with verbal ability. Behavioral genetic models of the data showed that environmental factors explained the majority of the variance in ToM performance in this sample. Shared environmental influences on verbal ability had a common impact on ToM and explained more than half the phenotypic correlation between these two skills. Possible underlying proximal mechanisms are discussed, including maternal speech and mind-mindedness, sibling interactions, and peer influences.

Children's abilities to understand others are dramatically transformed when they become aware that human actions are governed by mental states such as beliefs, desires, and intentions. A key stage in this acquisition of a theory of mind (ToM) is the recognition that beliefs involve representations of reality, and therefore can be mistaken (i.e., false). Although this understanding of false belief is typically achieved around 4 years of age (Wellman, Cross, & Watson, 2001; Wimmer & Perner, 1983), individual differences in false-belief comprehension can be striking and are associated with several important aspects of children's early social relationships. These include the frequency and sophistication of children's shared pretense (Astington & Jenkins, 1995; Hughes & Dunn, 1997; Taylor & Carlson, 1997; Youngblade & Dunn, 1995) and communication (Dunn & Cutting, 1999; Hughes & Dunn, 1998; Slomkowski & Dunn, 1996), as well as self-judgments and sensitivity to criticism (Cutting & Dunn, 2002; Dunn, 1995). Such associations suggest that ToM is pivotal to young children's social lives. What then accounts for individual differences in ToM?

Accelerated acquisition of a ToM has been reported for children from larger families (e.g., Lewis, Freeman, Kyriakidou, Maridaki-Kassotaki, & Berridge, 1996; Perner, Ruffman, & Leekam, 1994; Peterson, 2001; Ruffman, Perner, Naito, Parkin, & Clements, 1998). Cultural differences can also be found in both children's (Vinden, 1996) and adults' (Lillard, 1998) concepts of mind. Together with reports of delayed ToM among deaf children of hearing parents (Peterson & Siegal, 1995; Russell et al., 1998; Woolfe, Want, & Siegal, 2002), these findings highlight environmental influences on ToM acquisition. However, there is also evidence for genetic influences on ToM. For example, children with autism, a highly heritable disorder (Bailey, Palferman, Heavey, & Le Couteur, 1998), show profound impairments in ToM, and similar though more subtle impairments have also been found in girls with the chromosomal disorder Turner's syndrome (Skuse et al., 1997). Yet, extrapolating from pathology to normal individual variation may well be unwarranted, and direct assessment of genetic influences in a sample of typically developing children is therefore needed.

The most widely used genetically sensitive design is the twin study, which hinges on comparisons between monozygotic (MZ) twin pairs and dizygotic (DZ) twin pairs. This contrast provides what has been described as the perfect natural experiment (Martin, Boomsma, & Machin, 1997) in which to assess the bottom line of transmissible genetic effects.

We are grateful to the Environmental Risk (E-Risk) Longitudinal Twin Study mothers and fathers, the twins, and the twins' teachers for their participation. Our thanks to Michael Rutter and Robert Plomin and to members of the E-Risk team for their dedication, hard work, and insights. Terrie Moffitt is a Royal Society–Wolfson Research Merit Award holder. The E-Risk Study is funded by the Medical Research Council.

Correspondence concerning this article should be addressed to Claire Hughes, Centre for Family Research, University of Cambridge, Free School Lane, Cambridge CB2 3RF, UK. Electronic mail may be sent to ch288@cam.ac.uk.

Recently, Hughes and colleagues (Hughes & Cutting, 1999; Hughes & Plomin, 2000) reported on the first twin study of individual differences in ToM. In their study, 119 pairs of 42-month-old twins completed a battery of ToM tasks, and model-fitting analyses showed that 60% of the sample variance in ToM could be attributed to genetic factors. This finding supports theoretical accounts that highlight the innate nature of early ToM development (e.g., Baron-Cohen, 1995; Leslie, 1994). However, as the authors noted, because their study involved only a small (volunteer) sample, the results need to be replicated in a larger, more representative sample. Indeed, very large sample sizes may be needed to detect effects of shared environment (Martin, Eaves, Kearsley, & Davies, 1978).

Our first aim in the present study was to examine the relative contribution of genetic and environmental influences on individual differences in ToM in a nationally representative sample of 1,116 pairs of young twins. In our first set of results we set out the logic of the twin design before reporting both simple correlations for MZ and DZ twins' ToM performance and estimates from univariate statistical model-fitting analyses for the heritability of ToM—that is, the proportion of sample variance accounted for by genetic factors as well as by the magnitude of shared and nonshared environmental influences (in these analyses, genetic, shared, and nonshared environmental factors are termed A, C, and E, respectively).

Our second aim was to explore the possible overlap between influences on individual differences in ToM and in verbal ability. Numerous studies have reported a strong association between ToM and verbal ability (e.g., Astington & Jenkins, 1999; Cutting & Dunn, 1999; Farrar & Maag, 2002; Watson, Painter, & Bornstein, 2001). However, there is considerable controversy surrounding the nature and direction of causal relations between ToM and verbal ability. For example, some researchers have argued that specific linguistic structures (e.g., the embedded syntax of complementation; de Villiers, 2000; de Villiers & de Villiers, 2000; Tager-Flusberg, 2000) or functions (e.g., integration of different viewpoints; Nelson, *in press*) are necessary prerequisites for success on false-belief tasks. In contrast, others (e.g., Baldwin, Baird, Saylor, & Clark, 2001; Bloom & Tinker, 2001) have argued that fledgling ToM skills such as joint attention or intentional parsing provide the building blocks for language acquisition. These two views are not necessarily incompatible because the direction of influence is unlikely to be static within a developmental framework (Hughes, 2001; Tager-Flusberg, 2001). Longitudinal data are needed

to examine the direction of causal links between ToM and verbal ability; therefore, addressing this issue was beyond the scope of this cross-sectional study. Instead, we adopted a behavioral genetic model-fitting approach to bring a fresh perspective to the question of how ToM and verbal ability are related.

Specifically, because individual differences in verbal ability are heritable (Dale et al., 1998; Plomin, Owen, & McGuffin, 1994; Reznick, Corley, & Robinson, 1997), and genetic influences are rarely domain specific (Eley, 1997), it is possible that common genetic influences contribute to the association between individual differences in ToM and verbal ability. In support of this hypothesis, Hughes and Cutting (1999) found that just more than half of the phenotypic correlation between verbal ability and ToM in their sample of 42-month-old twins was mediated by common genetic influences. Alternatively, it may be that common environmental influences are of central importance in explaining the covariance between individual differences in ToM and verbal ability. For example, maternal speech is significantly associated with both language development (Snow, 1999) and success on false-belief tasks (Charman, Ruffman, & Clements, 2002; Ruffman, Perner, & Parkin, 1999; Vinden, 2001). However, this association may be at least partly genetically mediated. In the present study we used the genetically sensitive twin design and applied bivariate statistical model-fitting analyses to assess the relative magnitude of common genetic and environmental influences on ToM and verbal ability in 60-month-olds.

In sum, this investigation is the first large-scale population-based twin study of individual differences in ToM, and it has two main aims. Our first aim was to establish the relative significance of genetic, shared, and nonshared environmental influences on individual differences in ToM. Our second aim was to examine the contributions of common genetic and environmental influences to the relationship between individual differences in ToM and verbal ability.

Method

Participants

Participants are members of the Environmental Risk (E-Risk) Longitudinal Twin Study, which investigates how genetic and environmental factors shape children's development. The E-Risk sampling frame was based on two consecutive birth cohorts (1994 and 1995) in a birth register of twins born in

England and Wales (Trouton, Spinath, & Plomin, 2002). Of the 15,906 twin pairs born in these 2 years, 71% joined the register. Our sampling frame excluded opposite-sex twin pairs and began with the 73% of register families who had same-sex twins.

The E-Risk Study sought a sample size of 1,100 families to allow for attrition in future years of the longitudinal study while retaining statistical power. An initial list of families was drawn from the register to target for home visits, including 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 (a) to replace high-risk families who were selectively lost to the register via nonresponse and (b) to ensure sufficient base rates of problem behaviors given the low base rates expected for 60-month-olds. Early first childbearing was used as the risk-stratification variable because it was recorded 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 behaviors (Maynard, 1997; Moffitt & E-Risk Team, 2002). The sampling strategy resulted in a final sample in which two thirds of E-Risk Study mothers accurately represent all mothers in the general population (ages 15–48) in England and Wales in 1994 to 1995. The other one third of E-Risk 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 provide unbiased statistical estimates that can be generalized to the population, the data reported in this article were corrected with weighting to represent the proportion of young mothers in the United Kingdom (Bennett, Jarvis, Rowlands, Singleton, & Haselden, 1996).

Of the 1,203 eligible families, 1,116 (93%) participated in home-visit assessments. Zygosity was determined using a standard zygosity questionnaire that has been shown to have 95% accuracy (Price et al., 2000). Ambiguous cases were zygosity-typed using DNA. The sample includes 56% MZ and 44% DZ twin pairs. Sex is evenly distributed within zygosity (49% male). Data were collected within 2 months of the twins' fifth birthday (M age = 60 months, SD = 1.2 months). With parents' permission, research workers visited each home for 2.5 to 3 hr, in teams of two. While research worker one interviewed the mother, the other tested the twins in sequence in a different part of the house. Families were given gift vouchers for their participation, and children were given coloring books and stickers. All research workers had university degrees in behavi-

oral science, and experience in psychology, anthropology, or nursing. Each research worker completed a formal 15-day training program on the child assessment protocol to attain certification to a rigorous reliability standard.

Measures

Socioeconomic status (SES). To measure SES, we relied on information about the current (or last) occupation of mothers and their spouses or partners. This information was coded using the Office of Population Censuses and Surveys (1991) standard occupational classification system, which arranges occupational groups into six social classes: 1 = professional, 2 = managerial and technical, 3N = skilled (nonmanual), 3M = skilled (manual), 4 = partly skilled, and 5 = unskilled. Families were assigned the higher of the occupations held by the mother or her spouse or partner.

Verbal ability. The scale of our study required a measure of verbal ability that was simple and quick to administer; the vocabulary subtest from the Wechsler Preschool and Primary Scales of Intelligence–Revised (WPPSI–R; Wechsler, 1990) was chosen because it is widely used and well validated against more comprehensive assessments of verbal ability (Wechsler, Golombok, & Rust, 1992).

ToM. The ToM test questions were administered in a set order of increasing difficulty and were presented in a forced-choice format (or with a forced-choice prompt), accompanied by at least one control question to check story comprehension and recall. Children only received credit on a test question if they also passed the accompanying control question(s). Four standard ToM test questions tapped children's ability to attribute a first-order false belief to a story character (e.g., a mistaken belief about an object's identity or location). Four advanced ToM test questions tapped children's ability to make inferences from an attributed false belief (e.g., to predict how a character would feel as a result of his or her false belief), or to attribute a second-order false belief (i.e., a mistaken belief about a belief) to a story character (see the Appendix for task scripts and full procedural details).

The first task involved unexpected contents in a prototypical container (miniature pencils in a Smarties candy tube). To pass this task children were required to attribute a mistaken belief about the tube's contents to a puppet character. The second task was an object-transfer task involving two storybook characters (Sally and Andy). To pass this task, children were required to predict an action

based on an attributed false belief. Next were two belief-desire reasoning tasks that involved either a nice surprise or a nasty surprise. Each included both a standard ToM question (first-order false-belief prediction) and an advanced ToM question (emotion inference based on this attributed false belief). Finally, there were two second-order false-belief tasks that tapped children's ability to attribute to a story character a mistaken belief about another character's belief (the Granddad and Chocolates stories). For both the belief-desire reasoning tasks and the second-order false-belief tasks, children who responded correctly to advanced ToM questions were asked to justify their response, and they received a bonus point for each correct justification.

Summing across the eight test questions and four bonus questions gave a range of 0 to 12 for possible scores. In previous research we have shown that adopting an aggregate approach (Rushton, Brainerd, & Pressley, 1983), following standard psychometric testing practice in which each child's scores across multiple ToM tasks are summed, improves the reliability with which children's early mental-state awareness can be measured. The set of tasks used in this study showed acceptable 1-month test-retest reliability ($>.7$) with 5-year-old children across a wide range of abilities (Hughes et al., 2000). The internal consistency for this ToM scale was also acceptable (Cronbach's $\alpha = .64$).

Before conducting behavioral genetic analyses of individual differences in ToM, we tested for sex differences in mean levels of ToM and verbal ability by conducting separate ordinary least squares regression analyses, in which ToM and verbal ability

scores were each regressed on sex. These regression analyses were based on the sandwich or Huber-White variance estimator (Gould & Sribney, 1999), a method available in Stata 7.0 (StataCorp, 2001). Application of this technique addresses the assumption of independence of observations by penalizing estimated standard errors and thereby accounting for the dependence in the data due to analyzing sets of twins (i.e., 1,104 pairs or 2,208 children). With respect to verbal ability, the mean scores for boys ($M = 9.24$, $SD = 3.13$) and girls ($M = 9.27$, $SD = 2.94$) showed no significant difference, $b = -.03$, $SE = .17$, $p = .85$. With respect to ToM, however, mean scores were significantly lower for boys ($M = 4.63$, $SD = 3.25$) than for girls ($M = 5.16$, $SD = 3.33$), $b = -.52$, $SE = .19$, $p < .01$, although the effect size for this difference was small ($d = .16$).

Results

The 60-month-old children in the E-Risk Study showed marked individual differences in their ToM performance and in their verbal abilities, as shown by the descriptive statistics in Table 1. We used these data to address two questions. First, we examined genetic and environmental influences on ToM. Second, we examined the association between ToM and verbal ability, and tested what accounts for the association between these skills.

Genetic and Environmental Influences on ToM

Table 1 shows the within-pair twin correlations on ToM performance. We used these twin correlations

Table 1

Correlations Among Theory of Mind (ToM), Verbal Scores, and Socioeconomic Status (SES) for Monozygotic (MZ) and Dizygotic (DZ) Twins and Descriptive Statistics

		ToM (Twin 1)	ToM (Twin 2)	Verbal (Twin 1)	Verbal (Twin 2)	SES
MZ twins	ToM (Twin 1)	1.0				
	ToM (Twin 2)	.53***	1.0			
	Verbal (Twin 1)	.38***	.33***	1.0		
	Verbal (Twin 2)	.40***	.36***	.66***	1.0	
	SES	.29***	.24***	.27***	.25***	1.0
	<i>M (SD)</i>	4.74 (3.21)	4.88 (3.22)	9.15 (2.99)	9.14 (3.05)	4.87 (1.60)
DZ twins	ToM (Twin 1)	1.0				
	ToM (Twin 2)	.53***	1.0			
	Verbal (Twin 1)	.43***	.36***	1.0		
	Verbal (Twin 2)	.33***	.45***	.43***	1.0	
	SES	.30***	.30***	.30***	.26***	1.0
	<i>M (SD)</i>	5.23 (3.47)	4.79 (3.32)	9.46 (2.91)	9.30 (3.17)	5.00 (1.60)

Note. $N = 2,208$ twin children.
*** $p < .001$.

to estimate the contribution of genetic and environmental factors to individual differences in children's ToM. Using the classical twin design, population variation in ToM may be partialled into an additive genetic component and two types of environmental components by using the following logic. First, a genetic contribution to ToM is indicated when the similarity in ToM task scores is greater for MZ twins than for DZ twins. This inference is based on the fact that MZ twins share all their genes but DZ twins, like all siblings, share on average only half of the genes on which members of the human population can vary. Conversely, an environmental contribution to ToM is indicated if the similarity between MZ twins is less than twice the similarity between DZ twins. In model fitting, this component is called shared environmental variance, and it indexes environmental effects that can be detected because they have increased the similarity between siblings.

Twin studies also address the perennial question of why family members differ from one another (Plomin & Daniels, 1987). If MZ twins, despite sharing all their genes, are not perfectly identical for a phenotype such as ToM, this indicates that non-shared experiences, unique to each twin, reduce their similarity. In model fitting, this component is called child-specific, or nonshared, environmental variance. It indexes environmental effects that can be detected because they have created differences between siblings (phenotype measurement errors that are not shared by siblings can produce such effects too). For detailed explanations of the statistical methods that are applied to operationalize the logic behind behavioral genetic designs, see Plomin, DeFries, McClearn and McGuffin (2001).

We used maximum likelihood estimation techniques (Neale & Cardon, 1992) to test univariate models of children's ToM performance. These models decompose the variance in children's ToM into that which can be accounted for by latent additive genetic factors (A), shared environmental (C), and nonshared environmental factors (E, which also includes measurement error), hereafter called a univariate ACE model. Because the latent variables are unmeasured, they do not have a natural scale; instead, the variance is fixed at 1.0. To compare the fit of different models, we used two model-selection statistics. The first was the chi-square goodness-of-fit statistic. Large values indicate poor model fit to the observed covariance structure. When two models are nested (i.e., identical except for constraints placed on the submodel), the difference in fit between them can be evaluated with the chi-square difference, using as its degrees of freedom the degree-of-freedom differ-

Table 2
Parameter Estimates (Unsquarred) and Fit Statistics From Quantitative Genetic Models of Theory-of-Mind Performance

	A	C	E	χ^2	df	RMSEA (CI)
ACE	.26	.69*	.67*	3.88	3	.023 (0-.079)
CE		.73*	.69*	4.69	4	.018 (0-.069)

Note. A = additive genetic variance; C = common or shared environmental variance; E = nonshared environmental variance or error; RMSEA = root mean square error of approximation; CI = confidence interval.

* $p < .05$.

ence from the two models. When the chi-square difference is not statistically significant, the more parsimonious model is selected, as the test indicates that the constrained model fits equally well with the data. The second model-selection statistic was the root mean square error of approximation, which is an index of model discrepancy, per degree of freedom, from the observed covariance structure (MacCallum, Browne, & Sugawara, 1996). Values less than .05 indicate close fit and values less than .08 indicate fair fit to the data (Browne & Cudeck, 1993).

As Table 1 shows, MZ and DZ correlations for ToM were identical ($r = .53$), suggesting substantial shared environmental influence but negligible genetic influence on individual differences in ToM. Table 2 summarizes the goodness-of-fit statistics and parameter estimates from the quantitative genetic modeling of these data. The proportion of variance accounted for by the latent genetic and environmental factors can be calculated by squaring each of the parameter estimates. For example, genetic influences accounted for 7% of the variance in children's ToM (i.e., $.26 \times .26$). The strongest influences on individual differences in ToM were shared and nonshared environmental factors, which accounted for 48% and 45% of the variance, respectively.

Because genetic influences on ToM were nonsignificant in the full univariate model, we tested the fit of a more parsimonious model in which these genetic effects were hypothesized to be zero. The fit of the reduced model was not significantly different from the fit of the full model, $\chi^2_{diff}(1) = .81, ns$. Thus, genetic factors do not account for significant variation in 60-month-old children's ToM. Finally, we tested whether the results of the univariate ACE model differed for boys and girls by estimating one model in which the effects of the latent genetic and environmental factors were constrained to be the same for the two groups and a second model in which the effects of the latent genetic and environmental factors were allowed to differ for the two

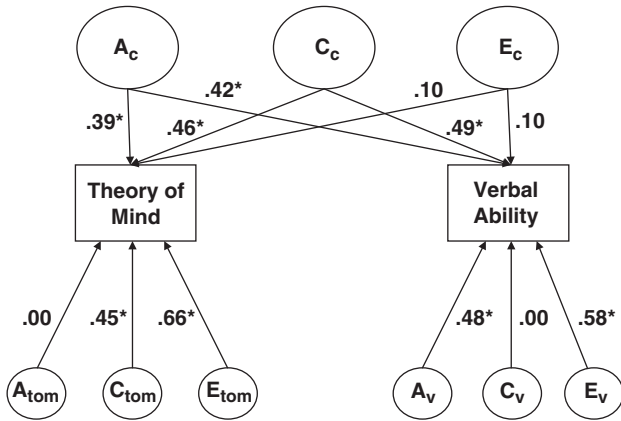


Figure 1. Variation in theory of mind (ToM) and verbal ability is accounted for by common genetic and shared environmental factors: Standardized parameter estimates for common factors model. A_{tom}, C_{tom}, E_{tom} = genetic, shared environmental and non-shared environmental influences that are specific to ToM; A_v, C_v, E_v = genetic, shared environmental and non-shared environmental influences that are specific to verbal ability; A_c, C_c and E_c = genetic, shared environmental and non-shared environmental influences that are common to ToM and verbal ability.

groups. The fit of the constrained unisex model was not significantly worse than the fit of the unconstrained model, $\chi^2_{diff}(3) = 2.05, ns$, indicating that the magnitude of genetic and environmental influences on ToM was similar for boys and girls.

What Accounts for the Correlation Between Children’s ToM and Their Verbal Ability?

Our second aim in this study was to examine the association between ToM and verbal ability from a behavioral genetic perspective. In the full sample of 2,208 sixty-month-old children, the *r* between ToM and verbal ability was .40 (*p* < .001). Using our genetically sensitive design, quantitative genetic anal-

yses can be applied to test the extent to which the correlation between ToM and verbal ability arises due to common genetic or environmental causes. A bivariate common factors model was specified to estimate genetic and environmental effects that are common to both ToM and verbal ability as well as genetic and environmental effects that are specific to ToM and verbal ability (Neale & Cardon, 1992). In bivariate twin analyses, MZ and DZ correlations were compared across traits (see Table 1); that is, ToM scores for one twin in each family (Twin 1) were correlated with verbal ability scores for their cotwins (Twin 2). Figure 1 shows the full bivariate ACE model used to investigate the extent to which individual differences in ToM and verbal ability scores could be attributed to common or distinct influences. The boxes indicate measured variables (ToM and verbal ability) and the circles indicate latent variables. Of these latent variables, those marked in capitals (i.e., A, C, E) have a common influence on ToM and verbal ability, and those marked in lower case letters (i.e., a_{tom}, a_v, c_{tom}, c_v, e_{tom}, e_v) have a unique influence on ToM or verbal ability, respectively.

The first row of Table 3 (see also Figure 1) presents the results of the common factors model. Individual differences in children’s ToM were predominantly accounted for by environmental factors (shared and nonshared) that were specific to ToM as well as genetic and shared environmental factors that were common to both ToM and verbal ability. Individual differences in verbal ability were accounted for by genetic and nonshared environmental factors that were specific to verbal ability as well as genetic and shared environmental factors that were common to both ToM and verbal ability. The only genetic factors that influenced 60-month-old children’s ToM were those that were shared with verbal ability, accounting for 15% (.39 × .39) of the variation in children’s

Table 3
Standardized Parameter Estimates for the Bivariate Common Factors ACE Model of the Association Between Theory of Mind (ToM) and Verbal Ability

	Specific ace						Common ACE ^a						χ^2	df	RMSEA (CI)
	ToM			Verbal ability			ToM			Verbal ability					
	a _{tom}	c _{tom}	e _{tom}	a _v	c _v	e _v	A	C	E	A	C	E			
ACE	0	.45*	.66*	.48*	0	.58*	.39*	.46*	.10	.42*	.49*	.10	16.61	11	.03 (0–.058)

Note. RMSEA = root mean square error of approximation; CI = confidence interval; A, C, and E = genetic, shared environmental, and nonshared environmental factors that are common to ToM and verbal ability; a_{tom}, c_{tom}, e_{tom} = genetic, shared environmental, and nonshared environmental factors that are specific to ToM; a_v, c_v, e_v = genetic, shared environmental, and nonshared environmental factors that are specific to verbal ability.

^aThe paths from the latent factors to the phenotypes were constrained to be equal but the standardized estimates differ because of variance differences in the phenotypes.

**p* < .05.

ToM and 18% (.42 × .42) of the variation in their verbal ability. Similarly, the only shared environmental factors to influence verbal ability were those that also influenced ToM, accounting for 21% (.46 × .46) of the variation in children's ToM and 24% (.49 × .49) of the variation in their verbal ability.

The best-fitting model excluded ToM-specific genetic influences, verbal-ability-specific effects of shared environment, and nonshared environmental factors that were common to ToM and verbal ability, $\chi^2_{\text{diff}}(3) = .49$, *ns*. In this model, ToM-specific shared and nonshared environmental factors accounted for 20% and 44% of the variation in ToM, respectively. Genetic factors and shared environmental factors that were common to both ToM and verbal ability accounted for an additional 16% and 20% of the variation in ToM, respectively. Verbal-ability-specific genetic and nonshared environmental factors accounted for 23% and 34% of the variation in verbal ability, respectively. Genetic factors and shared environmental factors that were common to both ToM and verbal ability accounted for an additional 19% and 24% of the variation in verbal ability, respectively.

The full bivariate model can be used to explain whether ToM and verbal ability are correlated because they have genetic influences in common, environmental influences in common, or both. To test these hypotheses, the correlation between ToM and verbal ability was decomposed into that which can be accounted for by the common genetic versus the common shared and nonshared environmental factors. To reproduce the Pearson correlation of .40 between ToM and verbal ability, the products of the paths that connect the two phenotypes via the latent genetic and environmental factors were summed (e.g., .39 × .42 + .46 × .49 + .10 × .10 = .40). Thus, the correlation between children's ToM and verbal ability was largely accounted for by genetic and shared environmental factors that were common to both. Genetic factors accounted for 41% of the correlation (e.g., .39 × .42 / .40) and shared environmental factors accounted for 56% of the correlation. The remainder was accounted for by nonshared environmental factors that were common to both ToM and verbal ability.

We tested whether boys and girls differed with respect to the association between ToM and verbal ability by comparing the fit of a model in which the effects of common and specific genetic and environmental latent factors were free to vary by gender with the fit of a model in which these parameter estimates were constrained to be equal for boys and girls. The Pearson correlation between verbal ability and ToM was .41 for girls and .40 for boys ($p < .001$).

The fit of the constrained unisex model was not significantly worse than the fit of the unconstrained model, $\chi^2_{\text{diff}}(9) = 7.37$, *ns*, indicating that common and specific genetic and environmental influences on ToM and verbal ability were similar in magnitude for boys and girls. This result supports and extends the finding from the univariate quantitative genetic analysis, showing similar genetic and environmental influences on ToM in boys and girls.

Discussion

Our first aim in this study was to establish the relative magnitude of genetic, shared, and nonshared environmental influences on individual differences in ToM. Our second aim was to examine common genetic and environmental influences on the association between ToM and verbal ability. We consider the study findings in relation to each aim and then present our conclusions and accompanying caveats.

Factors Underlying Individual Differences in ToM

In this large sample of 1,116 pairs of 60-month-old twins, 44% of the variation in ToM scores was accounted for by ToM-specific nonshared environmental influences, 20% by ToM-specific shared environmental influences, 21% by common shared environmental influences on ToM and verbal ability, and 15% by common genetic influences. Here we consider ToM-specific factors; common influences on ToM and verbal ability are considered in the next section.

Nonshared environment. The large proportion of variance in ToM explained by nonshared environmental factors fits with the consensus view from behavioral genetic studies that the nonshared environment is a key source of influence on behavioral development (Hetherington, Reiss, & Plomin, 1994). Nonshared influences include child-specific life events (e.g., accidents and illnesses) and, more important, siblings' contrasting relationships with parents, with each other, and with peers.

Differential parenting may be especially relevant for families with twins. Young children typically are keenly aware of minor injustices in family life, and though this often results in conflict, discussions about why one child has received special treatment might accelerate children's awareness of differences in points of view. In support of this hypothesis, longitudinal studies show that frequency of family talk about issues of conflict predicts children's later ToM (Dunn & Slomkowski, 1992). Our results indicate that this relation may differ for individual

children in the same family; exploring the mechanisms underlying these differences may be one fruitful avenue for future research. Within relationships, contrasting experiences are well recognized and include variation along several distinct dimensions, including dominance, competition, support, enjoyment of other's company, and understanding of other's thoughts and feelings (Hinde, 1997). Even though relationships between same-sex twins are likely to be more symmetrical than other sibling relationships, these contrasts remain important sources of nonshared environmental influence. This view is supported by findings from a recent study of 8- to 16-year-old same-sex twins, in which contrasting experiences of relationships within the family were partially attributable to genetic differences but also reflected the impact of nonshared environmental influences (Carbonneau, Eaves, Silberg, Simonoff, & Rutter, 2002).

Contrasting experiences outside the family home are typically highlighted by studies of older school-aged or adolescent samples (e.g., Crosnoe & Elder, 2002; Reiss et al., 1994). Yet, children in the United Kingdom begin school by age 5 and therefore their social horizons are broadened beyond the family to include teachers and peers. This increased diversity of the social environment, coupled with a reduced overlap between genetic and environmental influences (because children do not typically share genes with social partners outside the family), makes school and peer experiences an important source of nonshared environmental influence. Recent findings from a longitudinal study of early friendships highlight this point. Dunn and colleagues (Dunn, Cutting, & Fisher, 2002) interviewed 70 children about new friendships formed in the 1st year of school and found that social insight (rated from their interview responses) was independently predicted by their preschool sociocognitive skills (i.e., ToM and emotion understanding) and by their previous friends' preschool sociocognitive skills. Friendships with socially skilled peers, therefore, appear to enhance children's social understanding from a very early age.

Shared environment. With regard to ToM-specific effects of shared environment, interesting findings have emerged from recent attachment research. First, similar concordance in attachment ratings for MZ and DZ preschool twins suggests robust shared environmental influences on individual differences in attachment (O'Connor & Croft, 2001). Second, securely attached infants outperform other children as preschoolers on tests of ToM, even when effects of verbal ability are controlled, suggesting a specific relation between attachment and ToM (Fonagy,

Redfern, & Charman, 1997). Third, mind-mindedness (i.e., the propensity to attribute mental states to one's infant) has recently been highlighted as an important facet of maternal sensitivity that predicts both security of attachment and later ToM skills (Meins, Fernyhough, Fradley, & Tuckey, 2001; Meins et al., 2003; Meins et al., 2002). Together, these findings suggest that some specific shared environmental influence on ToM may be explained by individual differences in maternal mind-mindedness.

Explaining the Correlation Between ToM and Verbal Ability

Consistent with the literature, we found a strong phenotypic correlation between children's ToM and verbal ability; common effects of genetic and shared environmental influences accounted for most of this correlation. Indeed, the only genetic factors to contribute to individual differences in ToM were those that also influenced verbal ability. This contrasts with the strong, domain-specific genetic influence on ToM (and negligible shared environmental influence) reported by Hughes and Cutting (1999) but fits with the domain-general genetic effects typically reported in behavioral genetic studies (Eley, 1997).

Also of note is that shared environmental influences on verbal ability overlapped entirely with shared environmental influences on ToM. What kinds of family factors might constitute a shared environmental influence on both ToM and verbal ability? Two obvious candidates are other siblings and SES-related factors.

Siblings. Children with siblings are known to show accelerated success on ToM tasks (Lewis et al., 1996; Perner et al., 1994; Peterson, 2001; Ruffman et al., 1998), suggesting that distinctly child-like interactions between siblings (e.g., squabbling, teasing, games of make-believe) may stimulate children's awareness of others' thoughts and feelings. And although, as argued earlier, twins may have contrasting experiences within the twin relationship, their shared twin status, matching gender, and shared position with respect to sibling hierarchies may increase the similarity in their relations with other siblings.

That said, it is hard to judge whether sibling influences on ToM should be attenuated or strengthened in a twin sample. On the one hand, the universal presence of a same-sex, same-age sibling for all children in the sample is likely to attenuate the effect of family size on individual differences in ToM. On the other hand, evidence suggests that (a) it is only the presence of older siblings that fosters ToM

development (Ruffman et al., 1998), and (b) the link between ToM and family size is strongest for children who are linguistically less competent (Jenkins & Astington, 1996). Although adult twins do not differ in verbal ability from singletons (Posthuma, De Geus, Bleichrodt, & Boomsma, 2000), young twins show an initial lag (Rutter & Redshaw, 1991; Rutter, Thorpe, Greenwood, Northstone, & Golding, 2003). Together, these findings suggest that family size may matter just as much for twin samples as for singleton samples. Further work is needed to explore this hypothesis.

Research findings on the impact of siblings on verbal ability are equally complicated. For example, studies that include birth order as a predictor of developmental language impairments yield both positive (Stanton Chapman, Chapman, Bainbridge, & Scott, 2002) and negative (Hershberger, 1996; Paul & Fountain, 1999) findings. Moreover, although firstborns show an advantage on tests of lexical and grammatical skills, later-borns show superior conversational skills (Hoff-Ginsberg, 1998). Although in this study it was only possible to use vocabulary scores as an index of children's verbal abilities, future research should include a variety of measures to tap individual differences in lexical, syntactic, and pragmatic abilities, as sibling influences may vary significantly for these aspects of verbal ability.

SES-related factors. Numerous studies have reported SES-related contrasts in both ToM (Cole & Mitchell, 1998; Cutting & Dunn, 1999) and language use (e.g., Hoff, Laursen, & Tardif, 2002; Hoff-Ginsberg, 1998; Walker, Greenwood, Hart, & Carta, 1994), and these associations were found in this sample too (see Table 1). Researchers have also begun to integrate these two sets of findings. For example, in a recent cross-cultural study, Shatz and colleagues (Shatz, Diesendruck, Martinez Beck, & Akar, 2003) compared ToM performance in preschoolers who were native speakers of languages that either include explicit terms to refer to false belief (e.g., Turkish and Puerto Rican Spanish) or lack these explicit terms (e.g., Brazilian Portuguese and English). Their findings demonstrated a local effect of lexical explicitness but a more general and robust effect of SES on ToM.

It is possible that differences in the frequency, content, and form of parent-child conversations (e.g., talk about inner states, narrative talk about story characters) may be central to explaining SES-related contrasts in ToM (Cutting & Dunn, 1999). Maternal talkativeness is strongly related to family SES. Both overall talk and child-directed talk are generally reduced in low-SES families (Hart & Risley, 1995; Hoff et al., 2002). Maternal speech is also the best single predictor of a young child's vocabulary (Huttenloc-

her, Haight, Bryk, Seltzer, & Lyons, 1991) and has recently been shown to mediate the effect of SES on early vocabulary development (Hoff, 2003). Our findings show that shared environmental influences on vocabulary had a common influence on ToM, suggesting that maternal speech may also mediate the effect of SES on ToM; this hypothesis has yet to be tested. If, as is generally believed, acquiring a ToM is of fundamental importance for children's social adjustment, these findings regarding the influence of SES carry clear practical implications for intervention programs that aim to reduce the negative impact of adverse environments on young children.

To develop effective interventions (e.g., programs that boost child-centered family talk), practitioners need to understand why low-SES families may be less likely to engage in the kinds of sustained open-ended conversations or frequent book sharing that support developments in both verbal ability and ToM. Maternal mind-mindedness is one candidate worth considering, particularly as differences in mothers' willingness to attribute intentions to infants are known to relate to educational status (Reznick, 1999).

Motivational factors may also be important. For example, in a study of maternal views on child health, low-SES mothers placed more importance on physical aspects of health than on psychosocial issues (Cheng, Savageau, DeWitt, Bigelow, & Charney, 1996). If attitudinal differences underlie SES contrasts in children's linguistic environments, raising parental awareness of the advantage of family conversations for children's sociocognitive development may foster ToM development in children from low-SES families. Although this view may seem naively optimistic, it is bolstered by findings from recent intervention studies that demonstrate clear improvements in ToM task performance for children given multiple communicative cues (semantic, syntactic, and pragmatic) to mistaken beliefs (e.g., Guajardo & Watson, 2002; Lohmann & Tomasello, in press; Pillow, Mash, Aloian, & Hill, 2002).

Conclusions and Caveats

This large-scale study of 60-month-old twins revealed striking individual differences in ToM. In itself, this finding challenges early models of ToM development as largely complete by age 4. In addition, in contrast with a previous report of strong genetic influence on individual differences in ToM in 42-month-olds (Hughes & Cutting, 1999), the present behavioral genetic analyses pointed to significant environmental influences (both shared and non-shared) on individual differences in ToM. The

contrast between the two studies requires an explanation. The simplest of these is that Hughes and Cutting's (1999) study may simply have had insufficient power, as large sample sizes are needed to detect shared environmental effects. The E-Risk Study's stratified recruitment design resulted in a higher proportion of children from very low SES families than is found in most studies; this may also have increased the study's sensitivity to shared environmental effects because these typically appear especially powerful for children facing extreme or multiple disadvantage (Scarr, 1992). Alternatively, nativists may argue that the contrasting findings reflect age-related changes in the relative salience of genetic and environmental influences on individual differences in ToM. Although the two samples differed by only 18 months in mean age, this age difference may be significant, especially because it spans the transition to school, a normative life event associated with a rapid expansion in children's social horizons that might increase the salience of environmental influences on children's ToM skills. This kind of developmental shift is a feature of recent theoretical accounts (e.g., Meltzoff, Gopnik, & Repacholi, 1999; Tager-Flusberg, 2001) that adopt a hybrid approach in which innate mechanisms are held to govern very early milestones in ToM (e.g., joint attention skills, imitation, affective contact with others), whereas social environments are thought to influence later developmental milestones (e.g., concepts of representational mental states).

The present study has several strengths, including the size and diversity of the sample, the comprehensive battery of ToM tasks used, and the genetically sensitive design. Nevertheless, two caveats should be noted. The first concerns our use of a single vocabulary test as a measure of verbal ability. However, vocabulary shows substantial correlations with other aspects of verbal ability (Dale, Dionne, Eley, & Plomin, 2000). Moreover, findings from a wide variety of studies involving children from diverse family backgrounds indicate that individual differences in vocabulary and grammar are subject to the same set of forces (Snow, 1999). Thus, although vocabulary is not the key focus of theoretical accounts linking ToM and verbal ability, the use of vocabulary scores does not seriously affect the validity of our findings. The second caveat concerns the distinction between processes that underlie normative individual variation and those that underlie the deficits shown by atypical populations. Recent evidence suggests that genetic factors play a stronger role in accounting for atypical development than in explaining variation within the normal range (Dale

et al., 2000; Spinath, Harlaar, Ronald, & Plomin, 2004). Although our findings may not apply to atypical groups, they help elucidate the origins of normal variation in ToM in young school-age children. The next step for research in this field is to examine the consequences of these marked contrasts in ToM for young children's social lives.

References

- Astington, J., & Jenkins, J. (1995). Theory of mind development and social understanding. *Cognition and Emotion, 9*, 151–165.
- Astington, J., & Jenkins, J. (1999). A longitudinal study of the relation between language and theory-of-mind development. *Developmental Psychology, 35*, 1311–1320.
- Bailey, A., Palferman, S., Heavey, L., & LeCouteur, A. (1998). Autism: The phenotype in relatives. *Journal of Autism and Developmental Disorders, 28*, 381–404.
- Baldwin, D., Baird, J., Saylor, M., & Clark, M. (2001). Infants parse dynamic action. *Child Development, 72*, 708–717.
- Baron-Cohen, S. (1989). The autistic child's theory of mind: A case of specific developmental delay. *Journal of Child Psychology and Psychiatry, 30*, 285–297.
- Baron-Cohen, S. (1995). *Mindblindness: An essay on autism and theory of mind*. Cambridge, MA: MIT Press.
- Bennett, N., Jarvis, L., Rowlands, O., Singleton, N., & Haselden, L. (1996). *Living in Britain: Results from the General Household Survey*. London: Her Majesty's Stationery Office.
- Bloom, L., & Tinker, E. (2001). *The intentionality model and language acquisition* (Vol. 66) Boston: Blackwell.
- Browne, M., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. Bollen & J. Long (Eds.), *Testing structural equation models* (pp. 136–162). Newbury Park, CA: Sage.
- Carbonneau, R., Eaves, L., Silberg, J., Simonoff, E., & Rutter, M. (2002). Assessment of the within-family environment in twins: Absolute versus differential ratings, and relationship with conduct problems. *Journal of Child Psychology and Psychiatry, 43*, 1064–1074.
- Cheng, T., Savageau, J., DeWitt, T., Bigelow, C., & Charney, E. (1996). Expectations, goals, and perceived effectiveness of child health supervision: A study of mothers in a pediatric practice. *Clinical Pediatrics, 35*, 129–137.
- Cole, K., & Mitchell, P. (1998). Family background in relation to deceptive ability and understanding of the mind. *Social Development, 7*, 181–197.
- Crosnoe, R., & Elder, G. H., J. R. (2002). Adolescent twins and emotional distress: The interrelated influence of nonshared environment and social structure. *Child Development, 73*, 1761–1774.
- Cutting, A., & Dunn, J. (1999). Theory of mind, emotion understanding, language, and family background: Individual differences and interrelations. *Child Development, 70*, 853–865.

- Cutting, A., & Dunn, J. (2002). The cost of understanding other people: Social cognition predicts young children's sensitivity to criticism. *Journal of Child Psychology and Psychiatry*, *43*, 849–860.
- Dale, P. S., Dionne, G., Eley, T. C., & Plomin, R. (2000). Lexical and grammatical development: A behavioural genetic perspective. *Journal of Child Language*, *27*, 619–642.
- Dale, P., Simonoff, E., Bishop, D., Eley, T., Oliver, B., Price, T., et al. (1998). Genetic influence on language delay in two-year-old children. *Nature Neuroscience*, *1*, 324–328.
- de Villiers, J. (2000). Language and theory of mind: What are the developmental relationships. In S. Baron-Cohen, H. Tager-Flusberg, & D. Cohen (Eds.), *Understanding other minds: Perspectives from developmental cognitive neuroscience* (pp. 83–123). Oxford, England: Oxford University Press.
- de Villiers, J., & de Villiers, P. (2000). Linguistic determination and the understanding of false beliefs. In P. Mitchell & K. Riggs (Eds.), *Children's reasoning and the mind* (pp. 191–228). Hove, England: Psychology Press.
- Dunn, J. (1995). Children as psychologists: The later correlates of individual differences in understanding of emotions and other minds. *Cognition and Emotion*, *9*, 187–201.
- Dunn, J., & Cutting, A. (1999). Understanding others, and individual differences in friendship interactions in young children. *Social Development*, *8*, 201–219.
- Dunn, J., Cutting, A. L., & Fisher, N. (2002). Old friends, new friends: Predictors of children's perspective on their friends at school. *Child Development*, *73*, 621–635.
- Dunn, J., & Slomkowski, C. (1992). Conflict and the development of social understanding. In C. Shantz & W. Hartup (Eds.), *Conflict and development* (pp. 70–92). Cambridge, England: Cambridge University Press.
- Eley, T. C. (1997). General genes: A new theme in developmental psychopathology. *Current Directions in Psychological Science*, *6*, 90–95.
- Farrar, M. J., & Maag, L. (2002). Early language development and the emergence of a theory of mind. *First Language*, *22*, 197–213.
- Fonagy, P., Redfern, S., & Charman, T. (1997). The relationship between belief-desire reasoning and a projective measure of attachment security (SAT). *British Journal of Developmental Psychology*, *15*, 51–61.
- Gould, W., & Sribney, W. (1999). *Maximum likelihood estimation with STATA*. College Station, TX: Stata Press.
- Guajardo, N. R., & Watson, A. C. (2002). Narrative discourse and theory of mind development. *Journal of Genetic Psychology*, *163*, 305–325.
- Harris, P. L., Johnson, C., Hutton, D., Andrews, G., & Cooke, T. (1989). Young children's theory of mind and emotion. *Cognition and Emotion*, *3*, 379–400.
- Hart, B., & Risley, T. R. (1995). *Meaningful differences in the everyday experience of young American children*. Baltimore: Brookes.
- Hershberger, P. (1996). The relationship of medical and family characteristics to language development skills of low birthweight children at three years of age. *Infant Toddler Intervention*, *6*, 75–85.
- Hetherington, E. M., Reiss, D., & Plomin, R. (Eds.). (1994). *Separate social worlds of siblings: The impact of nonshared environment on development*. Hillsdale, NJ: Erlbaum.
- Hinde, R. (1997). *Relationships: A dialectical perspective*. Hove, England: Psychology Press.
- Hoff, E. (2003). The specificity of environmental influence: Socioeconomic status affects early vocabulary development via maternal speech. *Child Development*, *74*, 1368–1378.
- Hoff, E., Laursen, B., & Tardif, T. (2002). Socioeconomic status and parenting. In M. Bornstein (Ed.), *Handbook of parenting: Vol. 2. Biology and ecology of parenting* (2nd ed., pp. 231–252). Mahwah, NJ: Erlbaum.
- Hoff-Ginsberg, E. (1998). The relation of birth order and socioeconomic status to children's language experience and language development. *Applied Psycholinguistics*, *19*, 603–629.
- Hughes, C. (2001). From infancy to inferences: Current perspectives on intentionality. *Journal of Cognitive Development*, *2*, 221–240.
- Hughes, C., Adlam, A., Happe, F., Jackson, J., Taylor, A., & Caspi, A. (2000). Good test–retest reliability for standard and advanced false-belief tasks across a wide range of abilities. *Journal of Child Psychology and Psychiatry*, *41*, 483–490.
- Hughes, C., & Cutting, A. (1999). Nature, nurture and individual differences in early understanding of mind. *Psychological Science*, *10*, 429–432.
- Hughes, C., & Dunn, J. (1997). "Pretend you didn't know": Preschoolers' talk about mental states in pretend play. *Cognitive Development*, *12*, 477–499.
- Hughes, C., & Dunn, J. (1998). Understanding mind and emotion: Longitudinal associations with mental-state talk between young friends. *Developmental Psychology*, *34*, 1026–1037.
- Hughes, C., & Plomin, R. (2000). Individual differences in early understanding of mind: Genes, nonshared environment and modularity. In P. Carruthers & A. Chamberlain (Eds.), *Evolution and the human mind: Language, modularity and social cognition* (pp. 47–61). Cambridge, England: Cambridge University Press.
- Huttenlocher, J., Haight, W., Bryk, A., Seltzer, M., & Lyons, T. (1991). Early vocabulary growth: Relation to language input and gender. *Developmental Psychology*, *27*, 236–248.
- Jenkins, J. M., & Astington, J. W. (1996). Cognitive factors and family structure associated with theory of mind development in young children. *Developmental Psychology*, *32*, 70–78.
- Leslie, A. (1994). ToMM, ToBY and Agency: Core architecture and domain specificity. In L. Hirschfeld & S. Gelman (Eds.), *Mapping the mind: Domain specificity in cognition and culture* (pp. 119–148). Cambridge, England: Cambridge University Press.

- Lewis, C., Freeman, N. H., Kyriakidou, C., Maridaki-Kasotaki, K., & Berridge, D. M. (1996). Social influences on false belief access: Specific sibling influences or general apprenticeship? *Child Development*, *67*, 2930–2947.
- Lillard, A. (1998). Ethnopsychologies: Cultural variations in theories of mind. *Psychological Bulletin*, *123*, 3–32.
- Lohmann, H., & Tomasello, M. (in press). The role of language in the development of false-belief understanding: A training study. In J. W. Astington & J. Baird (Eds.), *Why language matters for theory of mind*. Oxford, England: Blackwell.
- MacCallum, R., Browne, M., & Sugawara, H. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychological Methods*, *1*, 130–149.
- Martin, N., Boomsma, D., & Machin, G. (1997). A twin-pronged attack on complex traits. *Nature Genetics*, *17*, 387–392.
- Martin, N., Eaves, L., Kearsley, M., & Davies, P. (1978). The power of the classical twin study. *Heredity*, *40*, 97–116.
- Maynard, R. (Ed.). (1997). *Kids having kids: Economic costs and social consequences of teen pregnancy*. Washington, DC: Urban Institute Press.
- Meins, E., Fernyhough, C., Fradley, E., & Tuckey, M. (2001). Rethinking maternal sensitivity: Mothers' comments on infants' mental processes predict security of attachment at 12 months. *Journal of Child Psychology and Psychiatry*, *42*, 637–648.
- Meins, E., Fernyhough, C., Wainwright, R., Clark Carter, D., Gupta, M. D., Fradley, E., & Tuckey, M. (2003). Pathways to understanding mind: Construct validity and predictive validity of maternal mind-mindedness. *Child Development*, *74*, 1194–1211.
- Meins, E., Fernyhough, C., Wainwright, R., Gupta, M. D., Fradley, E., & Tuckey, M. (2002). Maternal mind-mindedness and attachment security as predictors of theory of mind understanding. *Child Development*, *73*, 1715–1726.
- Meltzoff, A., Gopnik, A., & Repacholi, B. (1999). Toddlers' understanding of intentions, desires and emotions: Explorations of the dark ages. In P. Zelazo, J. Astington, & D. Olson (Eds.), *Developing theories of intention: Social understanding and self-control* (pp. 17–41). Mahwah, NJ: Erlbaum.
- Moffitt, T., & E-Risk Team (2002). Contemporary teen aged mothers in Britain. *Journal of Child Psychology and Psychiatry*, *43*, 727–742.
- Neale, M., & Cardon, L. (1992). *Methodology for genetic studies of twins and families*. Dordrecht, Netherlands: Kluwer.
- Nelson, K. (in press). Language pathways into the community of minds. In J. Astington & J. Baird (Eds.), *Why language matters for theory of mind*. Oxford, England: Blackwell.
- O'Connor, T. G., & Croft, C. M. (2001). A twin study of attachment in preschool children. *Child Development*, *72*, 1501–1511.
- Office of Population Censuses and Surveys. (1991). *Standard Occupational Classification*. London: Her Majesty's Stationery Office.
- Paul, R., & Fountain, R. (1999). Predicting outcomes of early expressive language delay. *Infant Toddler Intervention*, *9*, 123–135.
- Perner, J., Ruffman, T., & Leekam, S. R. (1994). Theory of mind is contagious: You catch it from your sibs. *Child Development*, *65*, 1228–1238.
- Perner, J., & Wimmer, H. (1985). "John thinks that Mary thinks that . . .": Attribution of second-order beliefs by 5- to 10-year-old children. *Journal of Experimental Child Psychology*, *39*, 437–471.
- Peterson, C. (2001). Influence of siblings' perspectives on theory of mind. *Cognitive Development*, *15*, 435–455.
- Peterson, C., & Siegal, M. (1995). Deafness, conversation and theory of mind. *Journal of Child Psychology and Psychiatry*, *36*, 459–474.
- Pillow, B. H., Mash, C., Aloian, S., & Hill, V. (2002). Facilitating children's understanding of misinterpretation: Explanatory efforts and improvements in perspective taking. *Journal of Genetic Psychology*, *163*, 133–148.
- Plomin, R., & Daniels, D. (1987). Why are children in the same family so different from one another? *Behavioural and Brain Sciences*, *10*, 1–60.
- Plomin, R., DeFries, J., McClearn, G., & McGuffin, P. (2001). *Behavioral genetics* (4th ed.). New York: Worth.
- Plomin, R., Owen, M. J., & McGuffin, P. (1994). The genetic basis of complex human behaviours. *Science*, *264*, 1733–1739.
- Posthuma, D., De Geus, E., Bleichrodt, N., & Boomsma, D. (2000). Twin-singleton differences in intelligence. *Twin Research*, *3*, 83–87.
- Price, T., Freeman, B., Craig, I., Petrill, S., Ebersole, L., & Plomin, R. (2000). Infant zygosity can be assigned by parental report questionnaire data. *Twin Research*, *3*, 129–133.
- Reiss, D., Plomin, R., Hetherington, E. M., Howe, G. W., Rovine, M. J., Tryon, A., et al. (1994). The separate worlds of teenage siblings: An introduction to the study of the nonshared environment and adolescent development. In E. M. Hetherington & D. Reiss (Eds.), *Separate social worlds of siblings: The impact of nonshared environment on development* (pp. 63–109). Hillsdale, NJ: Erlbaum.
- Reznick, J. (1999). Influences on maternal attribution of infant intentionality. In P. D. Zelazo, J. W. Astington, & D. R. Olson (Eds.), *Developing theories of intention: Social understanding and self control* (pp. 243–268). Mahwah, NJ: Erlbaum.
- Reznick, J., Corley, R., & Robinson, J. (1997). A longitudinal twin study of intelligence in the second year. *Monographs of the Society for Research in Child Development*, *62* (Serial No. 249).
- Ruffman, T., Perner, J., Naito, M., Parkin, L., & Clements, W. (1998). Older but not younger siblings facilitate false belief understanding. *Developmental Psychology*, *34*, 161–174.
- Ruffman, T., Perner, J., & Parkin, L. (1999). How parenting style affects false belief understanding. *Social Development*, *8*, 395–411.

- Ruffman, T., Slade, L., Rowlandson, K., Rumsey, C., & Garnham, A. (2003). How language relates to belief, desire, and emotion understanding. *Cognitive Development, 18*, 139–158.
- Rushton, J. P., Brainerd, C. J., & Pressley, M. (1983). Behavioral development and construct validity: The principle of aggregation. *Psychological Bulletin, 94*, 18–38.
- Russell, P. A., Hosie, J. A., Gray, C. D., Scott, C., Hunter, N., Banks, J. S., et al. (1998). The development of theory of mind in deaf children. *Journal of Child Psychology and Psychiatry, 39*, 903–910.
- Rutter, M., & Redshaw, J. (1991). Growing up as a twin: Twin-singleton differences in psychological development. *Journal of Child Psychology and Psychiatry, 32*, 885–895.
- 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. (1992). Developmental theories for the 1990s: Development and individual differences. *Child Development, 63*, 1–19.
- Shatz, M., Diesendruck, G., Martinez Beck, I., & Akar, D. (2003). The influence of language and socioeconomic status on children's understanding of false belief. *Developmental Psychology, 39*, 717–729.
- Skuse, D. H., James, R. S., Bishop, D. V. M., Coppins, B., Dalton, P., Aamodt-Leeper, G., et al. (1997). Evidence from Turner's syndrome of an imprinted X-linked locus affecting cognitive function. *Nature, 387*, 705–708.
- Slomkowski, C., & Dunn, J. (1996). Young children's understanding of other people's beliefs and feelings and their connected communication with friends. *Developmental Psychology, 32*, 442–447.
- Snow, C. E. (1999). Social perspectives on the emergence of language. In B. MacWhinney (Ed.), *The emergence of Language* (pp. 257–276). Mahwah, NJ: Erlbaum.
- Spinath, F., Harlaar, N., Ronald, A., & Plomin, R. (2004). Substantial genetic influence on mild mental impairment in early childhood. *American Journal of Mental Retardation, 109*, 34–43.
- Stanton Chapman, T. L., Chapman, D. A., Bainbridge, N. L., & Scott, K. G. (2002). Identification of early risk factors for language impairment. *Research in Developmental Disabilities, 23*, 390–405.
- StataCorp. (2001). *Stata statistical software: Release 7.0*. College Station, TX: Author.
- Sullivan, K., Zaitchik, D., & Tager-Flusberg, H. (1994). Preschoolers can attribute second-order beliefs. *Developmental Psychology, 30*, 395–402.
- Tager-Flusberg, H. (2000). Language and understanding minds: Connections in autism. In S. Baron-Cohen, H. Tager-Flusberg, & D. Cohen (Eds.), *Understanding other minds: Perspectives from developmental cognitive neuroscience* (pp. 124–149). Oxford, England: Oxford University Press.
- Tager-Flusberg, H. (2001). A re-examination of the theory of mind hypothesis of autism. In J. Burack, T. Charman, N. Yirmiya, & P. Zelazo (Eds.), *The development of autism: Perspectives from theory and research* (pp. 173–194). Mahwah, NJ: Erlbaum.
- Taylor, M., & Carlson, S. (1997). The relation between individual differences in fantasy and theory of mind. *Child Development, 68*, 436–455.
- 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.
- Vinden, P. (1996). Junín Quecha children's understanding of mind. *Child Development, 67*, 1707–1716.
- Vinden, P. (2001). Parenting attitudes and children's understanding of mind. *A comparison of Korean American and Anglo-American families. Cognitive Development, 16*, 793–809.
- Walker, D., Greenwood, C., Hart, B., & Carta, J. (1994). Prediction of school outcomes based on early language production and socioeconomic factors. *Child Development, 65*, 606–621.
- Watson, A. C., Painter, K. M., & Bornstein, M. H. (2001). Longitudinal relations between 2-year-olds' language and 4-year-olds' theory of mind. *Journal of Cognition and Development, 2*, 449–457.
- Wechsler, D. (1990). *Wechsler Preschool and Primary Scale of Intelligence—Revised*. London: Psychological Corporation.
- Wechsler, D., Golombok, S., & Rust, J. (1992). *WISC—III UK Wechsler Intelligence Scale for Children—3rd Edition UK Manual*. Sidcup, England: Psychological Corporation.
- Wellman, H., Cross, D., & Watson, J. (2001). Meta-analysis of theory of mind development: The truth about false belief. *Child Development, 72*, 655–684.
- Wimmer, H., & Hartl, M. (1991). Against the Cartesian view on mind: Young children's difficulty with own false beliefs. *British Journal of Developmental Psychology, 9*, 125–138.
- Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition, 13*, 103–128.
- Woolfe, T., Want, S. C., & Siegal, M. (2002). Signposts to development: Theory of mind in deaf children. *Child Development, 73*, 768–778.
- Youngblade, L. M., & Dunn, J. (1995). Individual differences in young children's pretend play with mother and sibling: Links to relationships and understanding of other people's feelings and beliefs. *Child Development, 66*, 1472–1492.

Appendix

Unexpected-Contents First-Order False-Belief Task (Wimmer & Hartl, 1991)

This task involved a prototypical box (Smarties box/matchbox) in which was placed some unexpected contents (miniature colored pencils/Lego bricks). At the start of the

story, the child was shown the prototypical box and asked what he or she thought was inside. Having given the expected response (“sweets” or “matches”), the child was shown and asked to name the actual contents of the box. (All children gave the expected response to this question, and all were able to name the actual contents). The box was then closed, and the child was asked an own-belief question, “Before you looked inside, what did you think was in the box?” and a control question, “What is in the box really?” Next, the child was introduced to a puppet (Charlie) and asked an other-belief question, “What does Charlie think is in the box?” as well as a second control question, “What is in the box really?” To pass each test question, children were also required to respond correctly to the accompanying reality control question.

Unexpected-Location First-Order False-Belief Task (Wimmer & Perner, 1983)

This task involved four pictures. The first introduced the story characters, showing Andy with a bag and apple, and Sally with a box. The children were then told a story in which Andy puts his apple in his bag to keep it safe (Picture 2), but while he is outside playing Sally transfers the apple to her box (Picture 3), and then goes out to play. Next, Andy returns because he wants a bite of his apple (Picture 4). At this point, the child was asked the test question, “Where will Andy look for his apple?” If the child did not respond, a prompt was given, “In the bag or in the box?” All children were also asked a reality control question, “Where is the apple really?” and a memory control question, “Where was the apple first of all?” Children were only rated as successful on the test question if they responded correctly to both control questions.

Belief-Desire Reasoning Task (Harris, Johnson, Hutton, Andrews, & Cooke, 1989)

The two stories in this task involved either a mean surprise or a nice surprise, and the children were asked to predict an emotion from an attributed false belief. The mean surprise story involved two puppets (Larry Lion and Chris Crocodile), a miniature Coke can, a miniature milk carton, and the following script:

This is a story about two friends, Chris the Crocodile and Larry the Lion. Chris is a very naughty crocodile, and likes to play tricks on his friend Larry. Now, Larry really likes Coke, mmmm. In fact it’s his very favorite drink. Look! Here is Larry’s can of Coke. (Q1: How does Larry feel when he gets a can of Coke?). Larry doesn’t like any other drinks though and he really doesn’t like milk, yuck, yuck. Look here’s some milk. (Q2: How does Larry feel when he gets some milk?). One day, Larry went out for a walk, and naughty Chris decided to play a trick on his friend Larry. He poured out the Coke “Pssshhh!” and he instead he poured in some

milk “Glug-glug-glug.” Then he put the milk away, and went outside to watch Larry through the window. Now when Larry comes back from his walk, he’s really thirsty. He can see the can on the table, but he can’t see what’s inside the can. (Q3: When Larry first comes back from his walk, how does he feel—happy or not happy? Q4: Why does he feel happy? Q5: What does Larry think is in the can? Q6: What’s in the can really? Q7: How does Larry feel after he’s had a drink—happy or not happy? Q8: Why is he not happy?)

The questions were presented in a counterbalanced forced-choice format. Children who failed the emotion question but passed the false-belief question were given a second attempt at the emotion question. To pass the false-belief question (Q5), children also had to pass the reality control question (Q6). Children only passed the emotion test question (Q3) if they also passed the reality control (Q6) and all emotion contingency (Q1, Q2, Q7, and Q8) questions. Children who predicted the correct emotion were also asked to justify their choice (Q4).

The nice surprise story included the same questions in the same order but involved a story in which Freddy Frog wants to give his friend Peter Puppy a nice surprise. This nice surprise story involved swapping felt tip markers for crayons in a prototypical crayon box.

Unexpected-Location Second-Order False-Belief Task

Two stories were given in this task. The first followed the procedures for the memory-aid version of the second-order false-belief task developed by Perner and Wimmer (1985), as used by Baron-Cohen (1989). Props used for this task included: an A3 laminated picture card (with a house in the center, a back and front garden, a chair in the front garden, a park, and a beach), four plastic figures, a boy, a grandfather, a grandmother, and a pram with a baby. After introducing the children to each character, the researcher told the following story:

One day Granny said, “I’m going to take baby for a walk in the park, do you want to come with us Johnny?” It’s a hot day so Johnny said, “I’m too hot, I don’t want to go for a walk!” So Granny went off to the park with baby, while Johnny went to play in the back garden, and Granddad sat at the front of the house. A little later Granddad saw Granny coming back from the park. “Where are you going?” he said. Granny replied “The park was shut, so I’m going to take baby to the sea instead.” Granddad said, “OK, I’m going to have a little sleep.” Next, Granny and baby walked by the back garden. “Hello Granny, I’m up here!” waved Johnny from the tree. Granny told Johnny that she and baby were going to the seaside.

At this point children were asked two control questions: (Q1) “Does Granddad know that Granny talked to Johnny?” and (Q2) “Where are Granny and baby really?” The

task was discontinued if the child continued to fail either or both of these control questions, despite repeating the story. If the child passed these control questions, the story continued:

A little later, Johnny was bored, and decided to go and find Granny and baby. He ran back through the house and called out, "Granddad, I'm going off to play with Granny and baby."

Children were then asked the test question (Q3) "Where does Granddad think Johnny will go?" and a justification question (Q4) "Why does Granddad think Johnny will go there?" The task also included a reality control question (Q5) "Where are Granny and baby?" and a memory control question (Q6) "Where did Granny go with baby first of all?"

The second story, based on the simpler, second-order task developed by Sullivan and colleagues (Sullivan, Zaitchik, & Tager-Flusberg, 1994), was presented in a picture book format with the following script:

Picture 1: Granddad has given Mary and Simon some chocolate to share. "Go and put it away now children," says Granddad. "You can have some when Mum says so."
Picture 2: The children run into the kitchen and put the chocolate in the fridge, then they go out to play.

Picture 3: A little later, Simon comes in for a glass of water. He goes to the fridge and he sees the chocolate. He wants to keep the chocolate all for himself, so he takes the chocolate out of the fridge and puts it in his bag.

At this point children were asked two control questions: (Q1) "Where does Mary think the chocolate is?" and (Q2) "Where has Simon put the chocolate really?" (The story was discontinued if children failed either of these questions.)

Picture 4: Oh look! Mary is playing by the window; she can see everything that Simon is doing! She sees him put the chocolate in his bag! Simon is so busy hiding the chocolate he doesn't see Mary watching him! Later Mum calls Simon and Mary in for tea. She says they can have some of the chocolate. So, Simon and Mary come running into the kitchen.

Children are then asked the test question (Q3) "Where does Simon think Mary will look for the chocolate?" and a justification question (Q4) "Why does Simon think that?" The task also included two control questions: a reality control question (Q5) "Where is the chocolate really?" and a memory control question (Q6) "Where was the chocolate first of all?"