Family Correlates of False Belief Understanding in Early Childhood: A Meta-Analysis.

Rory T. Devine

Claire Hughes

University of Cambridge

Author Note.

Rory T. Devine and Claire Hughes, Centre for Family Research, Department of Psychology, University of Cambridge, Free School Lane, Cambridge, CB23RQ, United Kingdom. Rory T. Devine was funded by the Isaac Newton Trust, Cambridge and the UK Economic and Social Research Council (ES/JO21180/1). We would like to thank Dominic Kelly for his assistance with data collection. Correspondence regarding this article should be addressed to Rory T. Devine via electronic mail: rtd24@cam.ac.uk

Key Words: false belief; theory of mind; family influences; meta-analysis; mental-state talk; mind-mindedness; siblings; socio-economic status.
Abstract
This meta-analysis seeks to integrate findings from 25 years of research on family correlates of young children’s false belief understanding (FBU). Using data from 93 studies of 3- to 7-year-old children, we examined the correlations between FBU and four of the most widely-studied family factors: parental socio-economic status, number of siblings, parental mental-state talk and mind-mindedness. FBU exhibited modest associations with each family variable and these associations held even when individual differences in verbal ability were taken into account. Moderator analyses revealed key child-related factors (e.g., age, gender) as well as methodological factors that amplified or attenuated the relations between FBU and each family variable. Crucially, available longitudinal data highlight the importance of family factors in the development of FBU.
The capacity to attribute mental states such as desires and beliefs to others in order to explain, interpret or predict behavior, or ‘theory of mind’, has fascinated scholars in a wide range of disciplines for over three decades. A quarter of a century ago, Dunn and colleagues (1991) conducted a pioneering study of fifty 33-month-olds and their mothers, which showed that the frequency of maternal talk about feelings predicted individual differences in the children’s ability to explain a character’s mistaken beliefs. This seminal study led to a dramatic growth in research on potential family influences on individual differences in children’s developing understanding of mind (for recent reviews see: Hughes & Devine, 2015). In recent years researchers have capitalised upon meta-analyses to integrate the plethora of findings on theory of mind. Unlike traditional narrative reviews, meta-analyses are comprehensive and can provide researchers with an estimate of the magnitude, variability and overall statistical significance of a particular effect as well as a platform for theory development through the analysis of potential moderators (Ellis, 2010; Field & Gillett, 2010; Rosenthal, 1984). In an era of scientific research plagued by scandals concerning psychology’s ‘replication crisis’ and claims about selective reporting, meta-analysis provides a useful tool to assess the weight of evidence supporting a particular claim (e.g., Simmonsohn et al., 2014; Yong, 2012).

With regard to children’s false belief understanding (FBU - arguably the most intensely studied aspect of theory-of-mind development), meta-analyses have been used to great effect to examine: (1) the age at which children first attain FBU (Wellman et al., 2001); (2) the links between FBU and individual differences in language ability (Milligan et al., 2007) and executive function (Devine & Hughes, 2014); (3) the potential influence of culture on FBU (Liu et al., 2008); and (4) important social outcomes related to individual differences in FBU (Slaughter et al., 2015). Apart from providing evidence about the magnitude of
effects, these meta-analyses have been used to test key developmental accounts of theory of mind and identify important avenues for future research.

Since the pioneering study by Dunn et al. (1991), a wide variety of potential social correlates of FBU have been studied over the past two decades (Hughes & Devine, 2015). These include family socio-economic status (SES), family size (i.e., number of siblings), parental mind-mindedness, parental mental-state talk, parental sensitivity, parenting styles, attachment security, sibling relationship quality, and peer interactions. Nativist modularity theorists (e.g., German & Leslie, 2004) claim that such findings do not contradict their theoretical position because correlations between theory-of-mind task performance and various measures of social interaction can be explained by other factors. That is, task performance correlates with measures of social interaction because of the relations between domain-general processes and social interactions (German & Leslie, 2004).

In response to the growth of research on social correlates of theory of mind, ‘social’ accounts of theory-of-mind development have emerged. In particular, Carpendale and Lewis (2006) have outlined a social constructivist theory that is based on the ideas of Vygotsky, Piaget and Wittgenstein and that places children’s experiences of social interaction at the heart of understanding the development of theory of mind. In a similar vein, Nelson (2004, p.119) has proposed that children enter “a community of minds” through discourse with parents and others about different perspectives and points of view. More recently, Heyes and Frith (2014) have likened the development of ‘mind reading’ to the mastery of print reading to support their claim that theory-of-mind development occurs during two-way interactions with more skilled members of the culture. Each of these accounts suggests that social experiences play an important role in the development of theory of mind. If indeed social experiences have a specific role to play in the development of theory of mind then individual differences in children’s FBU should be correlated with a range of social factors and these
correlations should be independent of shared social influences on general cognitive development in, for example, language ability. Crucially social factors should exhibit developmental or longitudinal relations with FBU. Although the burgeoning interest in the social origins of children’s theory of mind has led to a number of narrative reviews on family correlates of theory-of-mind development (e.g., Hughes & Devine, 2015), to date no concerted effort has been made to integrate the substantial body of literature on this topic or to evaluate whether the evidence supports social accounts of theory-of-mind development.

Our primary goal in the current meta-analysis was to consider the four family variables that have been most widely studied in relation to individual differences in FBU: family socio-economic status (SES), family size, parental mind-mindedness and parental mental-state talk. Our aim was to assess the overall average magnitude of the correlations between each of these four family factors and children’s FBU. We focused specifically on how these factors contribute to preschoolers’ explicit FBU (the most widely studied index of theory of mind in the preschool years) to minimise heterogeneity in the methods and in our sample’s age. Explicit or representational FBU is typically measured using tasks that require children to attribute mistaken beliefs to a character or predict or explain a character’s behavior based on his or her mistaken beliefs about an object’s location, contents or identity. Meta-analyses have revealed that, despite the surface differences between the tasks used to index preschoolers’, the age at which children typically pass explicit false-belief tasks is unaffected by the type of false-belief task used (e.g., change of location versus unexpected contents) (Wellman et al., 2001). Moreover, different types of explicit false-belief tasks show similar strength correlations with measures of language (Milligan et al., 2007) and executive function (Devine & Hughes, 2014).

With regard to family correlates of theory of mind, we focused on those correlates that showed consistency in measurement across studies and allowed us to examine both distal
and proximal family influences on individual differences in preschoolers’ FBU. To this end we included two distal correlates, family SES and family size (i.e., number of siblings), and two proximal correlates, parental mind-mindedness and frequency of parents’ conversational references to mental states. Note that our initial searches revealed that while a number of additional potential family correlates of FBU had been examined within the literature (e.g., sibling relationship quality; attachment security; parenting style), there were too few studies with consistent measurement to enable these correlates to be included in the current meta-analysis. Below, we outline the literature on each of the four family correlates of FBU that provide the focus for the current meta-analysis.

**Socio-Economic Status and False Belief Understanding**

The impact of family SES on children’s cognitive development has received considerable attention within the scientific literature (e.g., Hackman & Farah, 2009). SES is a multi-faceted construct and in child development research it is often measured through parental or family income, occupation or education or through some combination of these measures (Bradley & Corwyn, 2002). For example, children from lower SES families lag behind those from higher SES families in in language ability (Hoff, 2006) and in executive function (Noble et al., 2005). As noted in the above summary of quantitative reviews in this field, both language and executive function are closely related to preschoolers’ performance on false-belief tasks, such that a similar SES-related contrast in FBU might be expected. Early studies of FBU, however, generally involved small samples of middle-class preschool children (and often adopted a rather crude pass / fail approach that relied on a single task rather than a coherent task battery). As a result, the potential importance of family background for children’s FBU was largely overlooked. The first goal of the present study was to examine the overall magnitude of the relation between parental SES and individual differences in FBU. Given the overlap between SES and language ability, we also sought to
examine whether SES was related to children’s FBU independently of individual differences in language ability. Given that SES is a multi-faceted construct that can be indexed in a variety of ways, one question to be addressed in the current meta-analysis is therefore whether particular markers of SES are especially salient.

**Family Size and False Belief Understanding**

The first evidence that sibling relationships might play an important role in the emergence of FBU came from a study by Dunn et al. (1991), who found that 33-month-olds who frequently engaged in co-operative play with and older sibling were more likely than their peers to succeed on a false-belief task administered 7-months later. Of course, even longitudinal correlations do not demonstrate causality; however, it seems reasonable to hypothesise that children with siblings would have more opportunities to learn about others’ minds through conversations and play. In the first test of this proposal, Perner, Ruffman and Leekam (1994) found that family size (as indexed by number of siblings) was a strong correlate of FBU (the presence of each sibling appeared to provide an advantage that was roughly comparable with 6 months in age). This finding however has sometimes been difficult to replicate (e.g., Cole & Mitchell, 1998) and subsequent researchers have begun to revise the so-called ‘siblings effect’ (Ruffman et al., 1998, p. 161) by focusing specifically on the influence of older siblings or ‘child-aged’ siblings (Peterson, 2000).

Ruffman et al. (1998) argued that children would learn more about mental states from their more skilled older siblings than from younger siblings. Extending this, Peterson (2000) hypothesized that sibling interactions would only benefit FBU within specific age boundaries. Specifically, interactions with infant siblings (<12 months) and teenage or adult siblings (aged over 12) might not provide the linguistic or social opportunities for preschool children to reflect on or reason about mental states. To date, the existing empirical evidence is rather contradictory and so the current meta-analysis provides a valuable opportunity to assess
whether the advantage in FBU associated with having a sibling varies in magnitude according to the age or birth order of siblings.

**Parental Mental-State Talk and False Belief Understanding**

According to some social accounts (Lewis & Carpendale, 2006; Heyes & Frith, 2014), children’s emerging understanding of mind is underpinned not by individualistic cognitive development (e.g., in the capacity for meta-representation) but rather by socially mediated processes, described by Nelson (2004, p.119) as entering a ‘community of minds’. This idea was also first examined in studies led by Dunn (e.g., Dunn et al., 1991) who proposed that it is through talking about mental states that thoughts, memories and beliefs are brought to a child’s attention. Despite the effort involved in transcribing and coding family conversations, numerous studies have now examined the relations between parental references to mental states and children’s FBU. That said, these studies differ from each other in several key respects, including observational setting and whether or not the analyses take account individual differences in overall rates of talk (i.e., parental verbosity). Moreover it is unclear whether the influence of parental talk on children’s verbal ability might account for relations between individual differences in parental mental-state talk on children’s FBU. The current meta-analysis provided a valuable opportunity to examine these possibilities.

**Parental Mind-Mindedness and False Belief Understanding**

Mind-mindedness is defined as the propensity to view the infant as a psychological agent (termed ‘mind-mindedness’) and appears to be an independent predictor of later success on tasks that test pre-schoolers’ understanding of mistaken beliefs and emotion understanding (Meins et al., 2002). From a theoretical perspective, by providing new perspectives on understanding attachment security and parental sensitivity (Meins et al., 2001), the construct of parental mind-mindedness has had a significant impact on research on social-cognitive development. A key question for the current meta-analysis concerned the
equivalence of results from studies adopting alternative approaches to measuring parental mind-mindedness. For example, observations of parent-infant interactions can be coded for both the frequency and appropriateness of mind-related comments (e.g., Meins et al., 2001). Alternatively, interviews in which parents describe their children can be coded for the proportion of mentalistic (as opposed to physical or behavioral) attributions offered (e.g., Meins, Fernyhough, Russell & Clark-Carter, 1998). Given that there is now a robust literature demonstrating the importance of child-driven effects on parents (e.g., Bell, 1968), it cannot be assumed that these two measures (i.e., talking with versus talking about a child) will show similarly strong associations with FBU. One way in which the current meta-analysis contributes to the field is by addressing this simple but important methodological question.

**Summary of Aims**

In an effort to integrate 25 years of research on the family correlates of children’s FBU, our primary goal was to establish the magnitude of the relations between FBU in the preschool years with two distal (i.e., SES and family size) and two proximal family variables (i.e., parental mind-mindedness and parental mental-state talk). In addition, we sought to investigate the independence of the relations between each of these four family variables and children’s FBU by examining the unique association between these variables when individual differences in children’s verbal ability were taken into account. Using moderator analyses, we also sought to examine relations between family variables and FBU in terms of their the consistency (e.g., do effect sizes vary according to study features such as year of publication or age-range of participants?) and specificity (e.g., does the total numbers of siblings or just the total number of child-aged siblings matter for FBU?; do different methods of measuring mind-mindedness give rise to similar or different correlations with FBU?). Finally, where possible, we sought to examine the developmental relations between each of the four family
variables and children’s FBU by integrating data from available longitudinal studies in which children’s FBU was assessed at more than one time point.

Method

Search Strategy and Inclusion Criteria

We conducted a systematic search of the literature available before spring 2015. To minimize the potential heterogeneity between studies we focused on studies of typically developing children that included at least one false-belief task. We used three primary electronic bibliographic databases to identify relevant published literature (i.e., Scopus, PsychInfo, Proquest Dissertation Abstracts International) and entered the following terms to search for theory of mind: ‘theory of mind’, ‘false belief’, ‘social understanding’, ‘mentalizing’, ‘mindreading’, ‘belief-desire’, ‘appearance reality’. To capture family correlates on children’s theory of mind we combined these search terms with the following terms: ‘maternal’, ‘family’, ‘parental’, ‘mother’, ‘sibling’, ‘mental-state talk’, ‘mind-mindedness’, ‘socio-economic’, ‘conversation’, ‘emotion’, ‘talk’, ‘attachment’. We followed up these searches by hand-searching the reference lists from included studies. Where authors had published more than one study, we contacted them via electronic mail to ascertain whether the author had any unpublished findings.

Together these searches yielded 890 separate abstracts (including unpublished dissertations and manuscripts). The abstracts were assessed to determine whether they met five key inclusion criteria: (1) the study was available in English; (2) the report contained the findings of an empirical study; (3) the sample included typically developing children; (4) the study included at least one false belief task; (5) the study included at least one measure of family influence. Using a sample of 100 abstracts, we achieved 91% agreement when screening studies for inclusion for further analysis. Following screening, 212 abstracts were selected for in-depth review. After in-depth analysis we extracted 93 relevant datasets from
98 manuscripts (i.e., publications, dissertations, unpublished articles). Given the diversity of topics studied under the umbrella of family correlates of FBU, we performed four separate meta-analyses to examine the relations between FBU and: (1) SES; (2) family size (i.e., number of siblings); (3) parental mental-state talk; (4) parental mind-mindedness.

**Coding Scheme**

Consistent with previous meta-analyses of FBU (e.g., Devine & Hughes, 2014; Milligan, Astington & Dack, 2007), we recorded several key study features: year of publication; country of origin; number of participants; mean age of participants; specific false belief tasks used; inclusion of control questions; use of aggregate false belief scores. We also recorded the aspect of the family influence measured (i.e., SES, number of siblings, mental-state talk, mind-mindedness) as well as the approach to coding and scoring the particular measure of family influence (e.g., interview versus observation; total number of siblings versus number of child-aged siblings). Finally, we noted the correlation coefficient for the relation between FBU and the family measure. When false-belief task batteries were used, the correlations between aggregate scores and the family measure were chosen over single false-belief tasks. When non-independent data (i.e., multiple publications on the same dataset) were reported, we took the average effect size. For longitudinal studies, we used the average effect size when calculating the mean cross-sectional effect size and both effect sizes (i.e., from the first and second time point) when examining the mean longitudinal effect size.

**Data Analysis**

To ensure that studies with larger samples were given greater value in the meta-analysis, we used an inverse-variance weighted random effects model whereby each effect size was weighted by \( n - 3 \) (Field & Gillett, 2010). We analysed effects of continuous and categorical moderators using inverse-variance weighted linear regression, using the \( Q \) statistic to analyse effects of moderators (Wilson, 2001). In addition to examining zero-order
correlations, where possible, we also sought to determine the independence of the association between each of the four family measures and FBU by analysing data on the partial correlations between family measures and FBU controlling for individual differences in verbal ability. In order to establish whether there were developmental relations between each of the four family factors and children’s FBU, where possible, we used data from longitudinal studies in which FBU was measured at more than one time point. We used an auto-regressive approach to estimate the longitudinal association between each family measure and FBU by calculating the partial correlation between each family measure and later FBU (i.e., Time 2) controlling for earlier FBU (i.e., Time 1) (e.g., Hertzog & Nesselroade, 2003; Menard, 2002). All analyses were performed on Fisher-transformed \( r \) values (\( Zr \)) but for ease of interpretation we report our results using Pearson’s \( r \) (and 95% confidence intervals). Using guidelines from Ellis (2010), modest effects ranged from .10 to < .30, moderate effects ranged from > .30 to < .50 and large effects were >.50. To assess the robustness of the meta-analytic findings, we calculated the ‘Fail Safe N’ (\( N_{FS} \)), that is, the number of null findings it would take to reduce the mean effect size to a non-significant level (Rosenthal, 1984). The \( N_{FS} \) was judged against a critical value of \( 5K + 10 \) (where \( K \) = the number of effect sizes). If the \( N_{FS} \) exceeded this value it was considered a robust finding (Rosenthal, 1984).

In recent years, theorists have challenged some of the assumptions of traditional fail safe calculations in response to evidence of selective publication practices and what has come to be termed as ‘\( p \)-hacking’, that is, the practice of running multiple statistical analyses to produce statistically significant results (Simmonsohn, Nelson & Simmons, 2014). To counter these potential difficulties, Simmonsohn et al. (2014) have proposed the use of \( p \)-curves to assess the evidential value of a set of research findings. That is, by plotting the distribution of statistically significant \( p \)-values (i.e., \( ps < .05 \)), it is possible to rule out selective reporting as an explanation for a set of research findings. Simmonsohn et al. (2014) have argued that: (1)
When effects are genuine, $p$-curves will be right skewed; (2) when effects do not exist, $p$-curves will be flattened; and (3) when effects are the result of $p$-hacking, $p$-curves will be left skewed. We plotted $p$-curves to evaluate the evidential value of research findings on each of the four correlates of FBU.

We examined the heterogeneity of effect sizes using both Cochran’s $Q$ and the $I^2$ statistic and classified the degree of heterogeneity using cut-offs recommended by Higgins et al. (2003). When the number of available studies was sufficiently large, we used weighted linear regression analyses to examine potential moderators of the relations between FBU and each of the family measures. In each of these models, we regressed the effect sizes for the correlation between FBU and the family variable of interest on continuous predictors such as the year of publication, the sample size and mean age of participants. To determine whether aggregate scores for FBU yielded larger effect sizes, we created a dummy variable coding those studies that did (1) or did not (0) use a false-belief task battery. This enabled us to examine whether potential moderator variables exerted independent effects on the strength of the association between FBU and aspects of the family environment. In addition to these variables, for each meta-regression analysis, we also created dummy variables for coding key study features. These are described in each of the relevant sections.

Results

False Belief Understanding and Family Socioeconomic Status

In our systematic review, we identified 50 effect sizes for the association between SES and FBU, based on data from 7,320 3- to 6-year-old children (51% male) living in 11 different countries (Table S1). The forest plot in Figure 2 shows the individual effect sizes (and 95% confidence intervals) for each study in the meta-analysis as well as the overall mean effect size. There was a modest but statistically significant association between SES and FBU, $r = .18$, 95%CI [.14, .22], $Z = 8.84$, $p < .001$. The $N_{FS}$ was 2,956 which greatly
exceeded the cut-off of 245. The $p$-curve for the studies indicating a statistically significant zero-order correlation between SES and FBU ($K = 21$) is shown in Figure 1A. The curve was skewed to the right, $Z = -7.1$, $p < .001$, indicating that the findings were unlikely to be the result of selective reporting. We calculated the mean partial correlation between SES and FBU controlling for individual differences in verbal ability using data from 30 studies ($N = 5,098$). The weighted mean partial correlation was attenuated but remained statistically significant, $r = .11$, 95%CI [.06, .16], $Z = 4.24$, $p < .0001$.

There was significant variation in the strength of the correlation between SES and FBU, $-.09 \geq r \leq .56$, $Q(49) = 109.22$, $p < .001$, and the $I^2$ indicated substantial heterogeneity, $I^2 = 55.14\%$. Participants were aged between 36.5 and 73.44 months, $M = 54.03$ months, $SD = 8.94$. The median sample size was 81 and sample sizes ranged from 24 to 2232, $M = 146.40$, $SD = 314.61$. The majority (40/50) of the data were published. Most of these studies (41/50) included a battery of false belief tasks and just under half (22/50) used a multi-measure aggregate of SES. We examined these potential sources of heterogeneity between study results using inverse variance weighted meta-regression (see Table S2). The overall model accounted for 44.76% of the variance in effect sizes, $Q(8, 41) = 48.88$, $p < .001$. Four key predictors of effect size emerged. The relations between SES and FBU were stronger in earlier publications, $\beta = -.37$, $p = .002$. Participant age also mattered with effect sizes being stronger for older children, $\beta = .40$, $p = .007$, and weaker in studies with a narrow age range of participants, $\beta = -.30$, $p = .04$. Stronger results emerged from studies that used aggregate as opposed to single measure indicators of SES, $\beta = .25$, $p = .04$. Using data from five longitudinal studies ($N = 732$) measuring FBU at more than one time point we calculated the mean longitudinal correlation with SES. There was a small but significant longitudinal correlation between SES and children’s later performance on measures of FBU even when earlier performance was taken into account, $r = .12$, 95%CI [.04, .19], $Z = 3.11$, $p < .01$. 
False Belief Understanding and Number of Siblings

We identified 45 separate studies with data on the correlation between FBU and either the total number of siblings (or family size) \( K = 22 \), the number of older siblings \( K = 12 \) or the number of child-aged siblings \( K = 11 \) (Table S3). These datasets came from a total of 4,996 3- to 7-year-olds (51% male) living in 14 different countries. The forest plot in Figure 3 shows the individual effect sizes (and 95% confidence intervals) for each study in the meta-analysis as well as the overall mean effect size. There was a modest but statistically significant association between family size and FBU, \( r = .14, 95\% CI [.08, .20], Z = 5.10, p < .001 \). Although the overall effect was modest, the \( N_{FS} \) was 1,344 which greatly exceeded the cut-off of 235. The \( p \)-curve for the studies indicating a statistically significant zero-order correlation between number of siblings and FBU \( K = 18 \) is shown in Figure 1B. The curve was significantly skewed to the right, \( Z = -6.53, p < .001 \), indicating that the findings were unlikely to be the result of selective reporting. Next we calculated the mean partial correlation between siblings and FBU controlling for verbal ability using data from 22 studies \( (N = 2,536) \). The weighted mean partial correlation was once again modest but statistically significant, \( r = .12, 95\% CI [.05, .19], Z = 3.26, p < .001 \).

There was substantial heterogeneity in the strength of the reported associations between family size and FBU, \(-.28 \geq r \leq .69, Q(44) = 156.63, p < .001, I^2 = 71.91\% \).

Participants were aged between 36.76 and 86.5 months, \( M = 54.55 \) months, \( SD = 9.06 \). The median sample size was 90 and sample sizes ranged from 33 to 385, \( M = 111.02, SD = 73.04 \). The majority of these studies (40/45) were published and most (35/45) measured theory of mind using a battery of false belief tasks. We examined potential sources of heterogeneity between study results using inverse variance weighted meta-regression (see Table S4). The overall model accounted for 32.03% of the variance in effect sizes, \( Q(9, 35) = 50.18, p < .001 \). Three features were significantly associated with the reported effect size. First, older studies...
were more likely to report stronger effect sizes, $\beta = -.39$, $p = .002$. Second, studies based on samples with more boys than girls reported larger effects, $\beta = .21$, $p = .015$. Third, consistent with the prediction made by Peterson (2000), the positive effect of siblings was stronger for studies that focused on child-aged siblings, $\beta = .29$, $p = .002$, than for studies that tallied either the total number of siblings or the number of older siblings. There were no significant differences in effect size between studies that measured number of older siblings versus the total number of siblings, $\beta = .03$, $p = .748$. Lastly, to investigate the longitudinal correlation between number of siblings and children’s FBU, we extracted data from five longitudinal studies ($N = 504$) in which children completed false-belief tasks at more than one time point. Controlling for individual differences in early FBU, there was a modest but significant longitudinal correlation between number of siblings and later FBU, $r = .16$, 95%CI [.07, .24], $Z = 3.55$, $p < .01$.

**False Belief Understanding and Parental Mental-State Talk**

The links between parental mental-state talk and children’s FBU have been examined in 28 separate studies, involving a total of 1,914 3- to 5-year-old children (50% male) from seven different countries (Table S5). The forest plot in Figure 4 shows the individual effect sizes (and 95% confidence intervals) for each study in the meta-analysis as well as the overall mean effect size. There was a modest but statistically significant association between parental mental-state talk and FBU, $r = .21$, 95%CI [.14, .27], $Z = 6.40$, $p < .001$. The $N_{FS}$ was 630 which greatly exceeded the cut-off of 145. The $p$-curve for the studies indicating a statistically significant association between parental mental-state talk and FBU ($K = 13$) is shown in Figure 1C. While visual inspection of this curve shows that there were several studies clustered around $p = .04$, the curve was significantly different from that expected if no effect existed (i.e., flattened distribution), $Z = -2.21$, $p = .013$. We extracted data from 12 studies ($N = 896$) to examine relations between parental mental state talk and children’s FBU.
when controlling for individual differences in verbal ability. The weighted mean partial correlation was modest but significant, $r = .19$, 95%CI [.13, .24], $Z = 6.16, p < .0001$.

There was significant heterogeneity in the strength of the reported associations between siblings and FBU, $- .18 \geq r \leq .48$, $Q(27) = 48.17, p < .001, I^2 = 43.94\%$. Participants were aged between 36.5 and 69.2 months, $M = 51.98$ months, $SD = 8.55$. The median sample size was 60 and sample sizes ranged from 24 to 263, $M = 68.35, SD = 44.91$. The majority of the studies (20/28) were published and most (23/28) used an aggregate measure of FBU based on scores from a task battery. With regard to the setting, 10 studies were based on recordings of talk during unstructured play or conversations, 16 studies were based on recordings of talk during a picture-book task and 2 were based on data from a questionnaire. Only 12 studies accounted for verbosity when coding parental mental-state talk.

We examined these potential sources of heterogeneity between study results using inverse variance weighted meta-regression (see Table S6). The overall model accounted for 58.45% of the variance in effect sizes, $Q(9, 18) = 28.15, p = .009$. The strength of the association between FBU and parental mental-state talk was significantly related to three study features. First, the year of publication was a strong predictor of effect size with older publications reporting larger effects, $\beta = -.43, p = .012$. Second, published studies had stronger effects than unpublished studies, $\beta = .61, p = .003$. Third, studies that calculated mental-state talk as a proportion of overall talk reported significantly lower effect sizes than those that did not, $\beta = -.66, p = .003$. Interestingly, effect sizes were unrelated to the setting task used to elicit mental-state talk, $\beta = .09, p = .559$. Finally, using data from six longitudinal studies ($N = 383$) in which FBU was measured at more than one time point, we estimated the mean longitudinal correlation between mental-state talk and later FBU controlling for earlier FBU. There was a moderate and significant developmental association between mental-state talk and later FBU, $r = .29$, 95%CI [.19, .38], $Z = 45.74, p < .0001$. 

17
False Belief Understanding and Mind-Mindedness

The relations between parental mind-mindedness and FBU were investigated in 14 separate studies involving 1,261 3- to 5-year-old children (49.5% male) from five different countries (Table S7). Participants were aged between 44.5 and 61.68 months, $M = 52.36$ months, $SD = 5.66$. The median sample size was 87 and sample sizes ranged from 33 to 164, $M = 90.07$, $SD = 44.60$. Of the 14 studies, all but 1 study accounted for overall parental verbosity when coding for mind-mindedness. Eight studies measured mind-mindedness using a parental interview, five coded mind-mindedness from observations of parent-child interaction and one study calculated mind-mindedness using both methods. We excluded this study from our moderator analyses due to non-independence of the effects.

The forest plot in Figure 5 shows the individual effect sizes (and 95% confidence intervals) for each study in the meta-analysis as well as the overall mean effect size. There was a modest but statistically significant association between parental mind-mindedness and FBU, $r = .16$, 95% CI [.09, .23], $Z = 4.67$, $p < .001$. The $N_{FS}$ was 153 which exceeded the cut-off of 80. The p-curve for the studies indicating a statistically significant association between mind-mindedness and false belief understanding ($K = 7$) is shown in Figure 1D. The curve was not significantly different from that expected if no effect existed, $Z = -0.29$, $p = .384$. In situations such as this, Simmonsohn et al. (2014) have proposed testing whether the p-curve is significantly flatter than a curve arising from studies powered at 33%. P-curves that are significantly flatter than this cut-off are interpreted as lacking evidential value because any effects are too small to be detected with existing sample sizes (Simmonsohn et al., 2014). Our findings revealed that the p-curve was not significantly flatter than that produced by studies powered at 33%, $Z = -1.22$, $p = .112$. Together these results suggest that the p-curve for the relations between mind-mindedness and FBU was inconclusive meaning that more studies are required to establish whether these findings have evidential value (Simmonsohn et al.,
Using data from 11 studies (N = 1,191) we calculated the mean partial correlation between parental mind-mindedness and children’s FBU controlling for individual differences in verbal ability. The correlation between FBU and parental mind-mindedness was modest but significant, \( r = .19, 95\% CI [.13, .24], Z = 6.16, p < .0001. \)

While effect sizes varied, \(- .05 \geq r \leq .35\), the heterogeneity in the strength of the reported associations between parental mind-mindedness and FBU was low and non-significant, \( Q(13) = 18.26, p = .15, I^2 = 28.81\% \). Given the lack of heterogeneity in effect sizes, we did not conduct formal moderator analyses. However, we did specify a regression model in which we examined whether the method used to measure mind-mindedness (i.e., interview versus observation) was related to the reported strength of the association between mind-mindedness and FBU. This model did not account for significant variance in effect sizes, \( R^2 = .02, Q(1, 11) = 0.40, p = .536 \), indicating that the method of measuring mind-mindedness did not influence the overall effect size for the correlation between mind-mindedness and FBU. Finally, to investigate the developmental relations between parental mind-mindedness and children’s FBU we collated data from longitudinal studies in which children’s performance on measures of FBU were assessed at more than one time point. Although there were 8 longitudinal studies investigating the relations between parental mind-mindedness and children’s FBU only 3 of these studies (\( N = 322 \)) included measures of FBU at more than one time point. The mean partial correlation between parental mind-mindedness and later FBU (controlling for earlier false-belief task performance) was modest and marginally significant, \( r = .12, 95\% CI [-.01, .23], Z = 1.92, p = .05. \)

**Follow-Up Analyses**

Given the availability of studies in which SES was measured alongside at least one other family variable we sought to examine whether the correlation between SES and FBU could be accounted for by one of the other family measures. We calculated the mean partial
correlation between SES and FBU controlling for family size ($K = 8, N = 768$), parental mental-state talk ($K = 7, N = 620$) and parental mind-mindedness ($K = 8, N = 890$). Although attenuated, the relation between FBU and SES remained significant even when controlling for family size, $r = .15$, 95%CI [.05, .24], $Z = 2.89$, $p < .01$, parental mental-state talk, $r = .10$, 95%CI [.03, .19], $Z = 4.31$, $p < .01$, and parental mind-mindedness, $r = .18$, 95%CI [.08, .29], $Z = 3.33$, $p < .001$. That is, the relations between SES and FBU could not be explained entirely by family size, parental mental-state talk or mind-mindedness.

Next we compared the strength of the correlation between each of the four correlates of FBU: SES, number of siblings, parental mental state talk and mind-mindedness. In order to avoid violating the assumption of independence required for meta-analysis we selected just one effect size from each study and so the results presented here should be treated with caution. To obtain a sufficient number of studies in each category we first began by selecting the two smallest categories of study (i.e., those examining parental mind-mindedness and mental-state talk). Next we extracted effect sizes from studies reporting on the correlation between siblings and FBU. Finally we included the remaining effect sizes for the correlation between SES and FBU. This analysis involved 14 studies of parental mind-mindedness, 27 studies of parental mental-state talk, 37 studies of siblings and 15 studies of SES. We examined the relative strength of the association between each of these family measures and children’s FBU using a $Q$ statistic ANOVA analogue (Wilson, 2001). There was no significant difference in the strength of the association between each of the four family variables and children’s FBU, $Q(3, 89) = 1.39$, $p = .71$. That is, the mean correlation between FBU and parental mind-mindedness, $r = .17$, 95%CI [.08, .26], $Z = 3.54$, $p < .001$, parental mental-state talk, $r = .20$, 95%CI [.14, .27], $Z = 5.67$, $p < .001$, number of siblings, $r = .16$, 95%CI [.11, .22], $Z = 5.89$, $p < .001$, and SES, $r = .21$, 95%CI [.12, .30], $Z = 4.58$, $p < .001$, was consistent in magnitude. Note however that while these mean effect sizes were similar to
those reported in the main analyses these results were based on a selection of the available studies.

**Discussion**

In our meta-analyses, we sought to integrate over 25 years of research on the family correlates of children’s FBU by (1) determining the strength and independence of the association between individual differences in young children’s FBU and four widely-studied measures of the family environment, (2) identifying potential moderators of any observed associations, and (3) examining the evidence for developmental relations between these variables. First we looked at two distal family factors: family SES and family size, as indexed by number of siblings. Our first set of analyses (based on data from over 7000 children aged between 3 and 6 years) revealed a modest correlation between family SES and individual differences in FBU. Although modest, the association held even when individual differences in verbal ability were taken into account. Our moderator analyses revealed that the correlation between family SES and children’s FBU was stronger for older children and weaker in more recent publications. Our second set of analyses focused on the association between FBU and family size. Data from almost 5000 3- to 5-year-olds revealed a modest correlation between these two constructs that remained significant even when individual differences in children’s language ability were taken into account. Our moderator analyses showed that the so-called sibling effect was stronger when the total number of ‘child-aged’ siblings as opposed to the total number of siblings or total number of older siblings was counted. Moreover, sibling effects on FBU were stronger for studies in which there were a greater number of boys than girls. Echoing the widely reported ‘decline effect’, that is, when initial large effects are followed by more attenuated effects, (e.g., de Bruin & Della Salla, 2015), SES and sibling effects were stronger in early studies than in more recent studies.
Next we looked at two proximal family factors. Our third set of analyses, which drew on data from just under 2000 3- to 5-year-old children, focused on the relations between parental mental-state talk and children’s FBU. Our results showed a modest link between parental mental-state talk and individual differences in FBU. Importantly, the correlation between parental mental-state talk and FBU was independent of any shared effects on children’s language ability. Moderator analyses revealed that the link between parental mental-state talk and children’s FBU was unrelated to whether the data were collected in structured (e.g., book reading) or unstructured (e.g., conversation) tasks. That said, the strength of association was weaker in studies where researchers reported the correlation between FBU and parental mental-state talk as a proportion of overall talk (i.e., studies that accounted for parental verbosity) as opposed to studies in which the simple tally of mental-state terms was used. Our fourth and final set of analyses focused on the association between parental mind-mindedness and 3- to 5-year-old children’s FBU and applied data from over 1200 children. These analyses revealed a modest (but statistically significant) correlation between parental mind-mindedness and children’s FBU even when individual differences in children’s language skills were accounted for. Importantly, these effects were unrelated to whether mind-mindedness was measured from observational or interview data.

In each of our four sets of analyses, the available longitudinal data suggested that there were developmental associations between each of the four family measures and FBU in early childhood. While these findings provide support for the view that children’s social environments might play a causal role in the development of FBU it is worth noting that these results should be treated with appropriate caution. Of the 93 studies included in this meta-analysis only 22 of these studies were longitudinal and only 13 of these included measures of FBU at more than one time point. Controlling for stability in the dependent variable (i.e., FBU) is an important step in auto-regressive analysis and in establishing whether there are
developmental relations between two variables (e.g., Hertzog & Nesselroade, 2003; Menard, 2002). Although we centered our analyses on these 13 studies, it is not clear whether these studies provide a representative sample of the available data.

**Theoretical Implications**

At first glance, the findings from our meta-analysis appear to support social accounts of theory-of-mind development (e.g., Heyes & Frith, 2014; Lewis & Carpendale, 2006). Specifically, these findings suggest that individual differences in young children’s FBU are significantly (but modestly) related to both distal and proximal family factors. Moreover, the potential general effects of family environments on children’s language skills do not appear to explain the links between children’s family environments and FBU. Finally, the extant longitudinal evidence (although limited) suggests that each of the four family variables included in this meta-analysis play a role in the development of children’s FBU. Together these findings challenge nativist accounts which have largely ignored the social context of theory-of-mind development.

That said, closer scrutiny of our findings reveals a more complex and nuanced picture that presents three key challenges to a purely social interpretation of the data. First, the links between FBU and each of the four variables studied here were modest in magnitude. Indeed, each of the four factors (i.e., family SES, number of siblings, parental mental-state talk and parental mind-mindedness) accounted for between just 2 and 4% of the variance in individual differences in FBU. To put this in perspective, the effect sizes reported in this meta-analysis were weaker than those reported for the relations between FBU and language (Milligan et al., 2007) or executive function (Devine & Hughes, 2014) but equivalent in magnitude to the correlation between theory of mind and peer acceptance (Slaughter et al., 2015). Second, although we identified a sizeable number of studies in which we could calculate partial correlations controlling for individual differences in children’s language ability, it is possible
that other potential confounding variables (e.g., executive function, genetic factors) might account for the reported associations. Third, the processes by which each of the four family variables independently contributes to individual differences in children’s FBU remain unclear. We will now consider each of these three points in more detail.

With regard to the first point, if social factors matter for theory-of-mind development why were the effects so modest? Findings on the social correlates of other neuro-cognitive abilities (e.g., working memory, executive function, general intelligence) in childhood show similarly modest associations with family factors such as parental SES (e.g., Noble et al., 2007; Von Strumm & Plomin, 2015). One possible explanation for the rather modest correlations reported here is that individual differences in children’s FBU may exhibit non-linear relations with family SES, family size, parental mental state talk and parental mind-mindedness. Much of the existing work on the social correlates of theory of mind is based on the assumption that social factors have a linear relationship with theory-of-mind development. Hertzman and Boyce (2010) have challenged this linear approach to understanding social causation when studying child development. Instead, they have argued that social conditions (such as family SES) can exert unpredictable and non-linear associations with developmental outcomes. It is possible that child factors reduce or amplify the links between family factors and children’s FBU. For example, our findings showed that the relations between family SES and FBU were stronger in older children than in younger children and that the ‘sibling effect’ was more pronounced in studies with fewer girls.

Second, it is often assumed that correlations between social factors (such as parental SES or parental behavior) and individual differences in children’s cognition provide causal evidence about the role of experience in cognitive development (Dale et al., 2015). However, when parents and children are drawn from the same family these supposed environmental effects are confounded with genetic effects. It is therefore entirely possible that the
correlations observed between each of the four family variables and children’s FBU could be
driven by ‘evocative’ gene-environment correlations (e.g., a child with enhanced social-
cognitive skills might stimulate his/her parents to provide a linguistically enriched social
environment) or ‘passive’ gene-environment correlation (e.g., correlations between parental
mental-state talk and children’s theory of mind might reflect the transmission of parents’
abilities to understand others’ mental states). Without genetically sensitive designs (such as
adoption studies), cross-lagged longitudinal research designs (in which child-driven effects
on parental mental-state talk or mind-mindedness can be tested) or training studies it will not
be possible to rule out these alternative explanations for the observed links between
children’s FBU and aspects of the family environment (Bryant, 1991). While adoption studies
have yet to be undertaken, cross-lagged longitudinal findings suggest that the link between
early parental mental-state talk and later FBU is greater than that between children’s early
FBU and later parental mental-state talk (e.g., Ruffman et al., 2002).

The third challenge is that social accounts do not adequately explain the processes by
which the plethora of distal and proximal family factors relate to the development of FBU or
indeed why there is such a wide range of family correlates of FBU. Assuming that genetic
factors were not responsible for the observed links between FBU and family SES, for
example, what processes might link this relatively distal factor with children’s understanding
of mental states? Our analyses go some way toward answering this question in that, even
when individual differences in family size, mental-state talk or parental mind-mindedness
were taken into account, FBU remained correlated with SES. That said, rather than exerting a
direct influence on FBU, it is possible that the links between family SES and children’s FBU
could be mediated by other aspects of parent-child interaction or indirectly through effects on
children’s executive function (e.g., Noble et al., 2007).
While our results contribute to our understanding of the relations between FBU and SES, using meta-analysis it was not possible to examine fully the overlap, independence or interplay of the four family factors. For instance, how might having a large number of child-aged siblings promote the development of children’s theory of mind? If the development of theory of mind hinges upon interactions with more experienced children and adults (e.g., Heyes & Frith, 2014), then interactions with younger (and less experienced) siblings should not confer a benefit. While numerous studies have sought to identify what combination of siblings might stimulate FBU, far fewer studies have attempted to examine the processes by which siblings help theory-of-mind development. One obvious possibility is that particular aspects of the sibling relationship might provide fertile ground for the growth of mental-state understanding. In line with this perspective, positive sibling relationships (as rated by parents) and low levels of conflict (observed during child-sibling interactions) were associated with advanced performance on measures of theory of mind (Cutting & Dunn, 2006; Hughes & Ensor, 2005). These findings require further replication in larger samples and in longitudinal designs.

With regard to parental mental-state talk and parental mind-mindedness, our results revealed that both factors showed similar strength associations with children’s FBU. To date, just one published study (Laranjo, Bernier, Meins & Carlson, 2014) has examined both constructs but it remains unclear whether these two measures are related to one another as the authors did not report on the association between these variables. One might assume that parental mind-mindedness might be related to parents’ use of mental-state language during parent-child conversations. However, recent findings indicate that an individual’s proclivity to describe others in mentalistic terms varies according to the relationship between that individual and the person being described (Meins et al., 2014). Thus rather than being a general tendency to use mental-state terms, parental mind-mindedness might reflect the
quality of the parent-child relationship. In summary, progress in social accounts of theory-of-
mind development requires closer attention to the relative independence or overlap between
these four widely studied family variables as predictors of children’s FBU and whether or not
these factors exert specific effects on theory of mind or general effects on cognition.

Methodological Implications

Our meta-analytic findings provide at least three useful methodological insights for
undertaking future research. First, measurement mattered across each of the four family
variables studied. That is, the ways in which researchers measured each of the family
correlates influenced the magnitude of the observed association between each correlate and
children’s FBU. Specifically, aggregate measures of family SES (integrating education,
income and/or occupational class) were stronger correlates of children’s FBU than were
single indicators. Likewise, the number of child-aged siblings was a stronger correlate of
children’s FBU than was the total number of siblings or total number of older siblings.
Moreover, studies of parental mental-state talk and mind-mindedness in which researchers
accounted for overall verbosity (e.g., by calculating mind-minded or mental terms as a
proportion of overall comments or talk) reported more modest correlations between these two
variables and children’s FBU. Consistency in how family variables are measured and
conceptualised will make future comparisons between research findings more easily
interpretable. In addition, standardization of methods across studies might reduce the ‘decline
effect’ by providing a more stable index of the true effect (de Bruin & Della Salla, 2015).

Second, although the longitudinal data appear to implicate family factors in the
development of FBU in early childhood, the preponderance of cross-sectional studies in this
field limits theoretical progress. Longitudinal and training studies will be needed to test
whether parental mental-state talk or mind-mindedness are necessary and sufficient to
promote the development of children’s theory of mind (Bryant, 1991). Third, while it was
reassuring that reported associations between FBU and each of the four family correlates were unlikely to be due to publication bias or \( p \)-hacking, many of the included studies were based upon very small sample sizes. Given the modest correlations between each of the four family variables and children’s FBU and the possibility that these factors may be related to FBU in non-linear ways, future studies would benefit from using larger sample sizes. Conducting large-scale studies on family correlates of FBU can be expensive (e.g., due to the cost of transcribing and manually coding the content of parent-child conversations) but new technologies could reduce the cost of future studies. Wearable recording devices capable of tallying the number of words spoken by parents and children and the total number of conversational turns have now been used to study young children’s exposure to child-directed language in naturalistic settings without the need for careful transcription or coding (Weisleder & Fernald, 2013). For researchers using transcripts of parent-child talk, computer software can now rapidly code and analyse parent-child talk for the overall frequency of mental-state terms, conversational turns and a range of other important linguistic features (e.g., Tausczik & Pennebaker, 2010). As software improves, it may soon be possible to use wearable recording devices to monitor mental-state talk in naturalistic settings.

**Caveats and Conclusions**

Meta-analysis is a relatively crude instrument and so a number of additional limitations (not already mentioned) should be acknowledged. In an effort to reduce heterogeneity, we focused our analyses on studies published in English using the most widespread index of theory of mind in early childhood: the false-belief task. Our findings were therefore restricted in three ways. First, it is unclear whether the four family variables examined here show culturally-universal links with children’s FBU. Second, the standard false-belief task has a limited developmental range and so our analyses are silent on whether family factors are related to individual differences in theory of mind beyond early childhood.
Recent findings suggest that family factors such as parental-mental state talk (Ensor, Devine, Marks & Hughes, 2014) and SES (Devine & Hughes, 2016) show significant associations with more ‘advanced’ theory-of-mind measures in middle childhood. Third, although popular, the validity of the standard or ‘direct’ false-belief task has been questioned (e.g., Bloom & German, 2000). The past decade has witnessed the growth of a range of minimally verbal indirect or ‘implicit’ false belief tasks that rely on children’s gaze rather than explicit verbal or manual responses (Baillargeon et al., 2010). There is considerable debate about whether these tasks measure the same or distinct abilities as traditional ‘direct’ false-belief tasks (e.g., Perner & Low, 2012). With one notable exception (Meristo et al., 2012), the literature remains relatively silent on the links between children’s social experiences and individual differences in performance on these tasks. Consequently, the findings from this meta-analysis shed light on a rather narrow aspect of children’s theory of mind.

In conducting this meta-analysis it was clear that the strongest available data on potential family correlates of FBU came from studies of SES and the weakest data came from studies of parental mind-mindedness. Specifically, the meta-analysis of the relations between SES and FBU was based on the largest number of studies encompassing the most children while the findings on mind-mindedness and FBU were based on the smallest number of studies with the fewest children. This is unsurprising given that SES is easily and inexpensively measured using questionnaires while measures of mind-mindedness (and mental-state talk) require more time-consuming and detailed analyses of parent-child interactions. That said, the differences in the availability of data from large samples of children should be kept in mind when interpreting the results.

Notwithstanding these and the other limitations noted earlier, our meta-analysis makes an important contribution to research on children’s theory of mind. Our review is the first to provide a systematic integration of research findings on family correlates of young
children’s FBU. Our study also provides the first attempt to assess the influence of selective publication practices on the state of findings in the field of theory-of-mind research. We have identified four family factors that show modest but statistically significant links with individual differences in FBU in early childhood. Supporting social accounts of theory-of-mind development, our results suggest that the correlations between these family factors and FBU are not accounted for by individual differences in children’s verbal ability. Moreover, the available longitudinal data suggest that each of these factors exhibit developmental associations with FBU in early childhood. Theoretical progress will require more ambitious study designs. Moreover, the modest effects suggest that factors outside the home environment such as peer relationships (e.g., Banerjee et al., 2011) or formal education (e.g., Wang, Devine, Wong & Hughes, 2016) might also play an important role in the development of theory of mind.
A. SES and False Belief Understanding

B. Siblings and False Belief Understanding

C. Mental-State Talk and False Belief Understanding

D. Mind-Mindedness and False Belief Understanding

Figure 1. P-Curves for the detection of publication bias in each of the four reported meta-analyses.
Figure 2. Forest plot showing the relations between family SES and children’s false belief understanding.

Note. Dashed line represents the null effect and solid line represents weighted mean effect size. Error bars show 95% confidence intervals.
Figure 3. Forest plot showing the relations between number of siblings and children’s false belief understanding.

Note. Dashed line represents the null effect and solid line represents weighted mean effect size. Error bars show 95% confidence intervals.
Figure 4. Forest plot showing the relations between parental mental-state talk and children’s false belief understanding.

Note. Dashed line represents the null effect and solid line represents weighted mean effect size. Error bars show 95% confidence intervals.
Figure 5. Forest plot showing the relations between parental mind-mindedness and children’s false belief understanding.

Note. Dashed line represents the null effect and solid line represents weighted mean effect size. Error bars show 95% confidence intervals.
References


German, T., & Leslie, A. (2004). No (social) construction without (meta-) representation: Modular mechanisms as a basis for the capacity to acquire an understanding of mind. *Behavioral and Brain Sciences, 27*, 106-107. DOI: 10.1017/S0140525X04330037


http://mason.gmu.edu/~dwilsonb/ma.html