Child Development, May/June 2002, Volume 73, Number 3, Pages 734–751

The Relation between Children's and Mothers' Mental State Language and Theory-of-Mind Understanding

Ted Ruffman, Lance Slade, and Elena Crowe

This study investigated the relation between mothers' utterances and theory of mind in a longitudinal study involving three time points over 1 year. Mothers were asked to describe some pictures to 82 children at all three time points. Mothers' use of mental state utterances in these descriptions at early time points was consistently correlated with later theory-of-mind understanding. This was true even when a number of potential mediators were accounted for, including children's own use of mental state language, their earlier theory-of-mind understanding, their language ability, their age, mothers' education, and other types of mother utterances. Mothers' mental state utterances seemed genuinely causal because early theory-of-mind ability was not related to later mother mental state utterances (i.e., it was not a reciprocal relation). Results also showed that children's desire talk preceded their talk about beliefs.

INTRODUCTION

The present study examined two main issues. First, it examined whether mothers' mental state language plays a causal role in the development of children's theory of mind. Philosophers (e.g., Churchland, 1988) and psychologists (Wellman, 1990) have argued that desires and beliefs are the two mental states that are central to understanding others' behavior and emotions. Research has shown that children come to understand how desires affect emotions and actions around 2 or 3 years of age, and how beliefs do so at about 4 years of age (for summaries, see Perner, 1991; Wellman, 1990), although there are debates about the precise age of onset (e.g., Lewis & Mitchell, 1994; Moore et al., 1995). This study explored whether mothers' language relates to children's subsequent performance on tasks that tap children's understanding of desires, emotions, and beliefs. Second, it examined whether children's utterances about desires precede their utterances about thoughts, and act as a training ground for theory of mind in general. Our interest was in theory of mind construed broadly as mental state knowledge, rather than whether this knowledge is theorylike (Gopnik, 1993; Perner, 1991; Wellman, 1990), or simulation based (Gordon, 1992; Harris, 1992).

There are a number of reasons for thinking that parents affect theory-of-mind understanding. These include the demonstrated influence of other family members, the things mothers say in their disciplinary efforts, the kind of talk they use in play and nondisciplinary situations, and the parenting practices associated with secure child attachment (these reasons are discussed further below). A number of studies have shown that children who have many siblings pass theory-of-mind tasks earlier (Jenkins & Astington, 1996; Perner, Ruffman, & Leekam, 1994; Ruffman, Perner, Naito, Parkin, & Clements, 1998). Likewise, greater time spent with older kin and older adults (Lewis, Freeman, Kyriakidou, Maridaki-Kassotaki, & Berridge, 1996) is associated with improved theory-of-mind performance. Because parents typically spend such a large amount of time with their children, it would be surprising if they too did not influence their children's theory of mind.

Among the studies that have directly examined mother effects, some have used a single time point. For instance, Ruffman, Perner, and Parkin (1999) found advanced false-belief understanding in children of mothers who claimed they focused on the mental states or feelings of others when their children had transgressed against those others. Meins, Fernyhough, Russell, and Clark-Carter (1998) studied mother effects directly using a longitudinal study. They found that security of attachment at 11 or 13 months of age was linked to a number of subsequent child and mother attributes. Mothers of securely attached children tended to describe their children using mental state terms when the children were 37 months, and were more sensitive to their children's difficulties or success when giving feedback on a task in which children constructed an object from smaller objects at 37 months. Crucially, securely attached children did better on false-belief tasks at 49 months of age. They also found that mothers' sensitivity and tendency to use mental attributes when describing their children at 37 months were related to chil-

dren's false-belief understanding at 49 months, and that mental attribute descriptions were related to belief-based emotion understanding at 61.5 months (see also Meins & Fernyhough, 1998). Meins et al. (1998) argued that maternal "mind-mindedness" (mothers' tendency to focus on their children's independent mental states) facilitates theory-of-mind understanding.

In another longitudinal study, Dunn, Brown, Slomkowski, Tesla, and Youngblade (1991) found that (1) children's talk about feelings (talk about desires or emotions) was related to performance on tasks that tapped affective understanding and belief understanding, (2) child–mother talk (child to mother + mother to child) correlated with both affective understanding and belief understanding, (3) children's talk about causes was linked to affective understanding, and (4) child–mother talk about causes (child to mother + mother to child) correlated with belief understanding. Similarly, Dunn, Brown, and Beardsall (1991) found that emotion understanding at age 6.5 correlated with both child-to-mother feeling state talk at 3 years of age, $r = .42$, and with mother-to-child feeling state talk, $r = .40$. Further, Dunn, Brown, and Beardsall (1991) found that causal language between child and mother that occurred within feeling state talk correlated with emotion understanding, and Dunn, Brown, Slomkowski, et al. (1991) found that causal language between child and mother correlated with subsequent false-belief understanding.

Despite this positive evidence, previous research has not provided strong grounds for inferring that mothers' interventions are causal. Single time point data (Ruffman et al., 1999) are not conclusive because it is impossible to assess the causal direction of a correlation (e.g., to assess whether mother interventions facilitate theory of mind, or theory of mind shapes mother interventions). Further, even with longitudinal designs, Dunn and colleagues (Dunn, Brown, & Beardsall, 1991; Dunn, Brown, Slomkowski, et al., 1991) acknowledge that one must be very cautious in inferring causality. For instance, Meins et al. (1998) and Dunn and colleagues (Dunn, Brown, & Beardsall, 1991; Dunn, Brown, Slomkowski, et al., 1991) used a single measure at Time 1 (e.g., mother interventions) and a second measure at Time 2 (e.g., theory of mind). There are two reasons why it is difficult to infer a unique causal role for mother utterances from such data. First, it is possible that mother interventions at Time 1 predict theory of mind at Time 2 indirectly through variance they share with theory-of-mind understanding at Time 1. That is, theory of mind at Time 1 would be expected to correlate with theory of mind at Time 2. If mother interventions and theory of mind at Time 1 correlate with one another, then early mother interventions might correlate with later theory of mind simply through the shared variance with early theory of mind, and not because they have a unique causal role in facilitating theory of mind.

Similarly, children's early social precocity might shape early mother interventions. One indicator of early social precocity might be children's use of mental state language (as well as their early theory of mind). Indeed, Dunn, Brown, Slomkowski, et al. (1991) combined child-to-mother talk with mother-to-child talk. They did not report mother-to-child talk on its own so it is unclear whether mothers' interventions facilitated theory of mind independent of children's initial skill level. This is particularly likely because both child-to-mother and mother-to-child talk correlated with social understanding in the study by Dunn, Brown, and Beardsall (1991). It is uncertain, then, whether mother-to-child talk when children were 3 years of age facilitated subsequent emotion understanding, or whether mothers talked more about mental states when their children were 3 years because they had picked up on their children's precocious emotion understanding (evidenced through advanced child mental state talk).

Consistent with this idea, Brown, Donelan-McCall, and Dunn (1996) found a relation between children's ability to predict false beliefs at 47 months of age and their mental state language at the same age. Again, mothers might have picked up on such language as a sign of social understanding and might have increased their own mental state language as a result. Yet it would not be that mothers' mental state language facilitated children's later theory of mind, but rather, that children's social precocity at an earlier time point was associated with an enhanced theory of mind at a later time point. For these reasons it is essential to partial out early theory-of-mind ability and children's early mental state language when considering whether early mother interventions facilitate later theory-of-mind ability.

In sum, stronger grounds for inferring causality could be provided by collecting data on theory of mind and mother interventions at each of two time points. A causal role for mother interventions would be supported if Time 1 mother utterances predicted Time 2 theory of mind, after partialing out Time 1 theory of mind, child mental state language, and child general language abilities. Further, a unique causal role for mother utterances would be supported if the reverse relation did not hold. The present study examined these issues by asking mothers to talk to their children about a set of pictures displaying everyday situations (e.g., a man and a girl at the seaside, two

children fighting over a toy, a man scolding a boy). Mothers' and children's utterances were then recorded, transcribed, and coded for different kinds of mental and non-mental state language (e.g., causal or fantasy language).

The present study had six main goals. First, mother intervention and theory-of-mind data were collected at each of three time points in an effort to provide converging evidence over different sets of time points. Second, previous research has paid only limited attention to the non-mental state language used by mothers, but it is very possible that it is something else that mothers say that is responsible for facilitating theory of mind. There are a number of categories of utterances that were clearly represented in the present data or that have been identified as facilitative in previous research. These included (1) general descriptions of the stimuli present in the pictures; (2) fact-based teaching; (3) talk about causes; (4) linking talk, in which mothers linked events in the pictures to their children's own life; and (5) elaborative thematic talk that went beyond the content of the pictures themselves. General descriptions and fact-based teaching might enhance children's general knowledge, language, and general inquisitiveness about events including behavior. Causal talk between mothers and their children has been found to correlate with belief and affective understanding (Dunn, Brown, & Beardsall, 1991; Dunn, Brown, Slomkowski et al., 1991). Research also indicates that mothers' conversations about past events in their children's life correlate with theory of mind (Welch-Ross, 1997). Thus, a correlation between mothers' linking to the past and theory of mind might be expected because the linking measure in the present study primarily involved talk about past events. Our interest in elaborative talk was motivated by findings of a link between pretense—another measure of fantasy—and theory of mind (Astington & Jenkins, 1995; Jenkins & Astington, 1996; Taylor & Carlson, 1997; Youngblade & Dunn, 1995). Mothers who expand on the visible content of pictures might stimulate general abilities to think about abstract nonvisible entities, including mental states. In sum, the present study allowed unique insight into whether these non-mental state utterances were as, or more, important to theory of mind than talk about mental states. This facilitation might take place directly or indirectly because non-mental state utterances enhance children's language, and language as measured on standardized tests is a known correlate of theory of mind (e.g., Astington & Jenkins, 1999; Charman & Shmueli-Goetz, 1998; Cutting & Dunn, 1999; Happé, 1995).

On a related note, a third aim was to examine children's language with a standardized test at each of the three time points to determine whether language ability mediated the relation between mothers' interventions and theory of mind. Dunn and colleagues (Dunn, Brown, & Beardsall, 1991; Dunn, Brown, Slomkowski et al., 1991) used mean linguistic utterance—the mean length of each child's 10 longest utterances—as a measure of language ability. Although mean linguistic utterance measures linguistic production, it does not directly tap linguistic understanding. In contrast to the robust correlations between standardized language tests and theory-of-mind, mean linguistic unit correlated significantly with only one of three theory-of-mind measures over the two studies of Dunn and colleagues (Dunn, Brown, & Beardsall, 1991; Dunn, Brown, Slomkowski et al., 1991). Thus, it is important to control for language ability using standardized language tests.

Fourth, this study collected data on child mental state utterances. Recall that previous research has established that early child mental state utterances correlate with later theory of mind; thus, it is possible that early child utterances mediate any relation between mother utterances and theory of mind.

Fifth, this study compared different types of mental state language. In particular, Bartsch and Wellman's (1995) claim that parents' early talk about desires is a higher correlate of children's subsequent talk about thoughts and knowledge than is parents' early talk about thoughts and knowledge was investigated. Bartsch and Wellman examined the desire, knowledge, and belief utterances of children from the CHILDES database. This is a computerized database in which children's utterances in everyday situations were recorded at various time points. Their analyses were based on between 4 and 10 children, depending on the specific analyses conducted and the amount of data present for each child. These authors argued that parents' early talk about desires is a training ground for children's understanding of mental life generally, including beliefs. Nevertheless, Bartsch and Wellman acknowledged that their findings could only be regarded as preliminary because (1) they were based on only 4 children, and (2) general cognitive ability was a potential confound. The present study examined these issues with 56 children, again taking into account children's performance on a standardized language measure.

A final aim of this study was to examine Bartsch and Wellman's (1995) claim that children's talk about desire precedes their talk about thoughts, and their related claim that children's lack of talk about thoughts was not due to an absence of parental talk about thoughts (because parental talk about thoughts

outnumbered children's talk). Again, Bartsch and Wellman acknowledged that their finding of earlier desire talk was based on only 10 children, and their finding of more parental talk about thoughts was based on only 4 children's parents. In an effort to provide converging evidence, similar analyses were conducted on 56 children in the present study.

At all time points, the choice of theory-of-mind tasks was motivated by a wish to provide tasks that tapped both desire and belief understanding, were appropriate for the age group tested, and would not overtax children's attention span. With this in mind, fewer tasks were given at Time 1 than at Time 3, and some tasks (emotion–situation) were only given at Time 1, assuming that children would be near ceiling subsequently. Likewise, children were asked justification questions and were given a "wicked desires" and ambiguity task only at Time 3, assuming that these questions and tasks would be too demanding at earlier time points.

METHOD

Children were tested at three time points, although some children dropped out over time. Time 2 occurred 5 months after Time 1, and Time 3 occurred 12 months after Time 1.

Participants

At Time 1 there were 82 children (age: $M = 3.01$ years, *range* = 2.18–4.04 years; 41 girls and 41 boys); at Time 2 there were 79 children (age: $M = 3.41$ years, *range* = 2.60–4.46 years; 40 girls and 39 boys); and at Time 3 there were 72 children (age: $M = 4.04$ years, *range* = 3.40–5.06 years; 36 girls and 36 boys). Not all children had complete data at each time point; therefore numbers of children examined in individual analyses sometimes varied. Mothers of the children responded to advertisements in newspapers, parent group newsletters, and a parenting magazine, and were paid £40 (Sterling) for completing the study. Children were from middle- and upper middle-class rural and urban areas in a White area of the United Kingdom.

The variation in age between children at each time point is an advantage in that it permitted claims about a broader range of children. Nevertheless, in many analyses age was partialed out to ensure that the findings were not due to younger or older children only. One finding that validated grouping children together was that there was no correlation between mother mental state utterances and age at any time point (see below).

Materials and Procedure

Table 1 provides a summary of the tasks given at each time point.

Language. The language test given at all three time points was the linguistic concepts subtest of the Clinical Evaluation of Language Fundamentals–Preschool Test (CELF). Examples of items are provided in Appendix A. At Times 1 and 2, this subtest was given in full using the CELF guidelines. At Time 3 the subtest was started one item below the first point at which the child had failed two of three items at Time 2. For instance, if a child had failed items 12 and 14 at Time 2, the subtest was started at item 11.

Picture task. At Time 1, mothers were given 10 photographs of people involved in common situations (e.g., a woman bathing a baby, a girl playing on a climbing frame, a man and a girl at the beach). These pictures were taken from the Thorpe Interac-

Table 1 Tasks Employed at Each Time Point

	Time 1	Time 2	Time 3
Language	CELF: linguistic concepts	CELF: linguistic concepts	CELF: linguistic concepts
Theory of mind			
Belief	False-belief translocation	False-belief translocation False-belief contents other	False-belief translocation Translocation justification False-belief contents other Contents other justification False-belief contents self Ambiguity
Desire	Desire–emotion Emotion–situations	Desire–emotion Desire–action	Desire–action Wicked desires
Picture task	10 pictures	13 emotion pictures	13 emotion pictures

Note: CELF = Clinical Evaluation of Language Fundamentals.

tion Measure (Thorpe, 1996). Mothers were given the photos and the experimenter said, "Can you look at these pictures with [child's name], like you would with, say, a book at bedtime or pictures in a magazine?" The dialogue between mothers and children was audiotaped and later transcribed. Coders worked from the written transcriptions but used the audiotapes to disambiguate utterances as needed. To ensure that the results were not a factor of a particular set of pictures, at Time 2 and again at Time 3, mothers were given 13 new photographs. The new pictures portrayed more emotionally charged or mentalistic situations (e.g., a "father" scolding his "son," a girl showing obvious signs of shyness, two people apparently thinking about their next move in a chess game). This picture set would potentially have been more sensitive to picking up differences in mental state utterances between mothers. Our interest was in whether an identical pattern of results would be obtained with the picture set used at Time 1 versus Times 2 and 3, given the potential difference in measurement sensitivity. This is important to demonstrate the generalizability of the results irrespective of picture set and measurement issues. For example, if mothers displayed a similar pattern of variation across these disparate picture sets, it would go some way toward showing that a stable and wide-ranging conversational style was being tapped, rather than a tendency that is tied to a particular type of context.

Theory of mind. The false-belief transfer task given at all time points (see Table 1) was based on the study by Wimmer and Perner (1983), the desire–emotion task was based on Wellman and Woolley (1990), and the emotion–situations task was based on Denham (1986). In the false-belief transfer task a story character, Katy, placed a cake (made of clay) in a red box and left to go and sleep. A second character, Sam, found the cake, ate some, and placed the remainder in a green box. Katy woke up and returned and children were asked the "belief" question ("Where will Katy look first?"), and three memory questions ("beginning": "Where was the cake in the beginning?", "now": "Where is the cake now?", and "watch": "Did Katy watch when Sam moved the cake?"). To be counted as correct on the "belief" question, children had to answer the three memory questions correctly. At Time 1, only children age 3 years and above were given the false-belief transfer task under the assumption that 2-year-olds would fail. The correlations between theory-of-mind understanding and mother utterances were tested first under this assumption, and second, using only the children who were actually administered the false-belief task (see Results section).

There were two versions of the desire–emotion task. In both versions children were told that the story character (David) liked one thing (e.g., horses) but not something else (e.g., cows). Children were then asked memory questions about what David likes, and if they failed, David's preferences were restated and children were asked the memory question again. If children failed a memory question twice they were counted as failing the task. Then, David looked in one location and either found a horse or a cow and children were asked, "How does David feel. Does he feel happy or does he feel sad?" Finally, children were again asked the memory questions about David's preferences (with the same procedure following incorrect answers), David searched in a second location and found the remaining animal, and children were again asked about his emotion. Children's answers were scored as correct only if they were right on both emotion questions (as well as the memory questions).

For the emotion–situations task, children were shown four different emotion faces (happy, sad, scared, and angry), and asked to identify each. The experimenter corrected children if necessary, and then demonstrated the procedure by placing the frightened face on a girl who was in a stereotypically frightening situation (confronted with a spider). Children were then given 14 stereotypical situations in a fixed order (see Appendix B). All situations were illustrated with line drawings and the experimenter pointed to the protagonists in turn and to the relevant action (if appropriate). For each situation, children were asked, "How does [name of character] feel? Put a face on her to show how she feels." This task was initially given to 22 adults whose ratings are listed in parentheses in Appendix B. For eight situations either or both of two emotions were accepted as correct because adults had named either emotion.

The false-belief contents task given at Times 2 and 3 was based on Perner, Leekam, and Wimmer (1987), and the desire–action task was based on Wellman and Bartsch (1988). In each of the two versions of the desire–action task a story character (Billy) wanted something (e.g., string) to take to a particular location (e.g., his nursery). The experimenter told children that the string was either in a box or in a bag. Billy looked in the box and either found the string or found nothing. Children were then asked the "action" question ("What does Billy do after he finds/doesn't find some string. Does he look in the bag or does he go to his nursery?"), the "want" question ("What did Billy want?"), and the "find" question ("What did Billy find?"). The want and find questions were control questions. To be scored as correct on the desire–action task children had to correctly answer all control ques-

tions and both "action" questions (when Billy found the string and when he found nothing).

In the false-belief contents task the experimenter introduced a doll (Julie) but then hid her in a bag. The experimenter then asked children to say what was in a crayons box, showed children that there was a key inside instead of crayons, checked whether children remembered what was inside, retrieved Julie and reminded the children that she hadn't seen inside the crayons box, and then asked children the "other" question, "When we first show Julie this box, before she looks inside, what will she say is in here?" (giving a forced choice if there was no answer). If children initially failed to acknowledge that there was a key in the box, the experimenter repeated the process of showing them what was inside. Children were counted as failing the task if they were incorrect on the question about the box's contents two times. At Time 2, children were just asked the "other" question. At Time 3, they were also asked a "self" question: "When I first showed you the box, all closed up like this, what did you think was in there then?"

At Time 3, children were also asked a justification question for both the contents "other" question ("Why will she think there are crayons inside?") and the false-belief translocation task ("Why will Katy look there?"). In the contents task, two children referred to the story character's lack of knowledge (e.g., "because her doesn't know") and were given half a point for their justification. Children who referred to the box's misleading appearance were given 1 point (e.g., "because it's got a picture of crayons"). In the translocation task, children who referred to where the story character originally put the cake, to her thinking that it had not moved, or to the character's lack of visual access were given 1 point (e.g., "because that's where she put it," "because she thinks it didn't move away," "'cuz she didn't see him"). One child who said "because her doesn't know" was given half a point on the assumption that the child meant the character didn't know about the object's transfer. These coding categories were based on studies by Wimmer and Mayringer (1998) and Clements, Rustin, and McCallum (2000).

The Time 3 wicked desires task was based on the study by Yuill, Perner, Pearson, Peerbhoy, and van den Ende (1996). A story character tried to throw a ball at a second character because he did not like him. In one story the second character caught the ball, and in the other story the ball hit the character. Children were then asked the test question: "How does the orange boy feel about that: happy or sad?" Next, children were asked two control questions: what did the first character want to do in the beginning, and why did he want to do that (in forced choice format, if necessary, e.g., "Did he like the green boy?"). Stories were illustrated by line drawings and children were scored as correct if they answered both control questions correctly and said the thrower would feel happy when he hit the second character and sad when he missed. A previous study by Yuill et al. (1996) found that 5-year-olds, but not 3- and 4-year-olds, showed this pattern.

The Time 3 ambiguity task was based on a study by Taylor (1988). Children were shown line drawings of two objects: first a flower and a rabbit, then an elephant and a giraffe. Each of the two objects was covered leaving only a small but visible bit that in the first case, made the flower and rabbit indistinguishable, and in the second case, made it possible still to identify the elephant and giraffe. Children were then asked whether a doll, that had not seen the full pictures, would know what each one was: "Does Mary know which one is the rabbit/elephant?" To be correct, children needed to pass both the ambiguous and the unambiguous condition.

Design

Tasks were given in a semirandom order as follows: the language test, one or more theory-of-mind tasks, the picture task, and any remaining theory-of-mind tasks.

Scoring

Theory of mind. At all three time points each task was given equal weighting. For instance, at Time 1 the 14 emotion–situation scenarios were scored as a proportion (out of one) and summed along with the false-belief transfer task (1 point) and the desire–emotion task (1 point) for a total of 3 points. At Time 2 the maximum theory-of-mind score was 4 points, and at Time 3 it was 8 points.

Language task. The raw number of correct items on the language task was used in the analyses.

Picture task. A summary of utterance categories, examples, and interrater reliabilities (Cohen's κs) is included in Table 2. Reliabilities were calculated for each codeable utterance (as designated by the primary coder), over both child and mother utterances together. A large majority (at least 80% at all time points) of the parents' utterances about desires, thoughts, and knowledge genuinely referred to mental states according to the criteria used by Bartsch and Wellman (1995) and described below. Although in some analyses genuine mental state utterances were examined, in others all mental utterances were grouped together. This was because (1) our primary concern was not in using mental state language to infer theory-of-mind

Table 2 Interrater Reliabilities for Mother and Child Utterances on the Picture Task

Category	Examples	Cohen's κ
Mental state utterances		
Desire	Want, like, love, hope, wish, dream, prefer, keen on	.94
Emotion	Happy, sad, unhappy, feel, cross, angry, grumpy	.96
Modulations of assertion	Might, maybe, perhaps, possibly, probably, could be, must, certainly, definitely, sure, guess, figure, reckon, certain, suppose, wonder, expect, curious, bet	.91
Think and know	Do you know what that is? She knows that's going to happen. They're thinking hard. Let me think. I think it's lovely.	.90
Other mental state	We went to France, remember? I understand that.	.92
Non-mental state utterances		
Descriptions	She's riding a bicycle. [person is riding a bicycle]	.85
Elaborations of a theme	What fish will he catch? [no fish visible]	.73
Causal talk	Why is he pointing? They have no clothes because they're in the water.	.70
Factual talk	A stethoscope is for listening to the heart.	.83
Links to child's life	We did that when we went to the beach.	.87
Don't know	I don't know.	.97
Physical state	Hurt, ill, in pain, sleepy, tired, hungry, thirsty, cry, smile, laugh, giggle	.86
Orienting utterances	Look, what's that?	.81
Repetitions of other	Mother repeats child's utterance.	.81
Self-repetitions	Mother repeats own utterance without a codeable intervening utterance.	.70

development in children, and (2) it is plausible that general use of mental state language might assist children if it can be understood as elaborating on behavior.

We were interested in the number of utterances that mothers and children made rather than the proportion (e.g., of all utterances). Potentially, each utterance has a direct bearing on children's theory of mind because it provides additional input, just as having more older siblings and spending more time with older individuals have been shown to facilitate theory of mind (Lewis et al., 1996; Ruffman et al., 1998). Two control measures were then employed to ensure that outliers were not having an undue effect on the data. First, the number of mental state utterances was broken into quartiles and z scores, and second, the number of mental state responses was compared to the number of other types of utterances to ensure that number of utterances was not simply picking up on parental involvement.

The broad mental state categories used by Bartsch and Wellman (1995) were also used in the present study. First, the cognitive terms "think" and "know" were coded together. Genuine mental state utterances included use of "think" as a mental activity (e.g., "They're thinking hard"), "think" used to refer to beliefs (e.g., "I think it's a cat," "Why do you think that?", "How do you think they did that?"), "think" used to refer to desires ("I think it's lovely," "They think that's good fun"), and rarely used contrastives (e.g., "I thought it was a cat, but I was wrong"). Think terms that could primarily mean "Yes" or "No" (e.g., "I think so"), and terms used for turn taking ("What do you think?") were coded as conversational rather than genuine uses.

"Know" terms coded as genuine mental utterances included those that referred to a lack of knowledge (e.g., "I don't know what it is," "I don't know how to do it," "I don't know whether it's a dog"), and those that questioned the source of knowledge (e.g., "How do you know that?"). "I don't know" responses (i.e., responses that consisted of only these three words and did not elaborate on what was unknown) were not coded as mental state terms because of their possible use to mean simply "I can't answer" (Bartsch & Wellman, 1995; Shatz, Wellman, & Silber, 1983).

Bartsch and Wellman's (1995) criteria were also followed when coding for desire terms. Terms used for social convention (e.g., "I don't care") were not counted as genuine mental utterances. Two other types excluded from the genuine category by Bartsch and Wellman—objectless statements of desire (e.g., "I wanna") and desire used to describe behavior (e.g., "I want a cookie," which might mean "Give me a cookie")—did not occur in the present sample. In the picture task, parents and children almost always described the de-

sires of the people in the pictures rather than their own desires. These descriptions were inferences based on story characters' expressions and actions, just the sort of elaborative description that might help children learn that there are mental states behind one's behavior. Hence, they were coded as genuine mental utterances. Parents and children usually used "want" when referring to desires, but also used "like," "love," "dream," "hope," "wish," "prefer," and "keen on."

Modulations of assertion included "might," "maybe," "perhaps," "possibly," "probably," "could be," "must," "certainly," "definitely," "sure," "guess," "figure," "reckon," "certain," "suppose," "wonder," "expect," "curious," and "bet." Context was used to determine whether an utterance was truly intended to modify certainty. For instance, mothers often wavered over an object in a picture and said, "I wonder what that is? It could be a cat." These terms express uncertainty or certainty and are correlates of belief understanding (Brown et al., 1996; Moore, Pure, & Furrow, 1990). Like belief, they focus on the subjective nature of mental states. "Think" and "know" could also be used to modify certainty, but following Bartsch and Wellman's (1985) criteria, were included in the "think and know" category.

The category "other mental state terms" described some sort of mental activity and included "remember," "understand," "forget," "remind," "realize," "idea," "consider," "have in mind," "daydream," "dream" (when asleep), "mean," "imagine" "wonder," and "expect." Note that "wonder" and "expect" could also be coded as modulations of assertion, and context was always used to assign the correct category. For instance, "I expect so" in response to a question about an object's identity might be treated as a modulation of assertion, whereas "He expects her to cry" might be coded as a mental state term rather than an expression of uncertainty.

The final mental state category was emotion terms, including "happy," "sad," "unhappy," "feel" (e.g., feel bad), "cross," "angry," "grumpy," "scared," "afraid," "disappointed," "worried," "upset," "surprised," "pleased," "enjoy," "excited," "fun," "interested," "frustrated," "missed," "annoyed," "hurtful," "bored," and "fed up." As in Shatz et al. (1983), desire (e.g., "want," "hope," "like") and emotion (e.g., "afraid," "happy," "sad") utterances were differentiated, in part, because emotion terms often have a more well-defined behavioral manifestation.

There were several categories of non-mental state utterances. The most common category was simple descriptions of a picture's contents (e.g., "She's riding a bicycle"). These comments added nothing more to the pictures, although they possibly enhanced interest and perhaps language. Physical state terms were also coded (Wellman, Harris, Banerjee & Sinclair, 1995) including "cry," "smile," "laugh," "giggle," "hurt," "in pain," "ill," "sleepy," "tired," "hungry," and "thirsty." Some of these utterances (e.g., "laugh") had strong links to emotions, but were coded separately because they described physical manifestations, whereas an emotion term such as "happy" could also refer to an internal experience. Because they do not refer to internal experiences, physical utterances would conceivably be less facilitory of theory of mind.

Causal utterances occurred when parents or children talked or asked questions about causes (e.g., "Why is he pointing?", "They have no clothes because they're in the water"). Thematic elaborations occurred when parents or children expanded on the visible content of a picture, for instance, saying "I wonder what fish he'll catch," when there was no fish visible. Links to the children's lives occurred when parents connected the pictorial information to their children's life (e.g., "We did that when we went to the beach"). Factual utterances involved some attempt to teach children general principles that did not include causal information (e.g., "A stethoscope is for listening to the heart"). Orienting responses included attempts by parents to focus children's attention on a picture (e.g., "Look, what's that?").

Utterances in which the mothers or children repeated themselves, with no codeable intervening utterances, were coded as self-repetitions rather than as another instance of the original utterance. Utterances in which mothers repeated their children's utterance (or vice versa) with no change in content, were coded as repetitions of the other. A statement could be coded in multiple ways. For instance, "You were crying because you felt sad" includes a causal statement (because), an emotion statement (sad), a physical state term (crying), and is also a link to the child's past (link). This statement therefore included four separate codeable utterances, and reliability between the two raters was calculated on all four types.

One coder used the transcriptions and audiotapes to code all utterances, and a second person coded 25% of the transcripts. The second coder was given examples of utterances but needed to rely heavily on context to categorize utterances. Second coding was established over all utterances (mother and child) because they were of an equivalent form and difficult to identify. Second coding was also established for individual categories to ensure that every category could be agreed on rather than achieving a good rating over all categories, with individual categories perhaps being suspect. Decisions about whether a new utterance was present were based on raters' ability to differen-

tiate each category of utterance. Kappas were all in a good range (see Table 2).

Maternal education. Mothers' education was coded on a 7-point scale as a measure of socioeconomic status (SES). Descriptions of this scoring with approximate North American equivalents was as follows: 0 = dropped out before age 16, 1 = Certificate of Secondary Education (high school up to age 16, with a focus on applied topics), 2 = General Certificate of Secondary Education or O-Levels (high school up to age 16, with a focus on academics), 3 = A-Levels (high school up to age 18, with a focus on academics), 4 = professional school (e.g., nursing, architecture), 5 = university undergraduate, and 6 = university postgraduate.

RESULTS

Table 3 provides a summary of the descriptive statistics for the theory-of-mind and language tasks, mothers' and children's mental state utterances, and maternal education. There were no floor or ceiling effects on any of the task totals. Mental state utterances were much more common for mothers than for children, giving the sense that mothers were scaffolding theory-of-mind language. One mother and her child were omitted because the mother's mental state utterances were over 3 *SD*s above the mean. Table 4 includes information about the frequency of occurrence of different mother utterances.

Over the three time points mothers were consistent in their tendency to be a high or low user of mental state language. The correlations between mental state usage were: for Times 1 and 2, $r(59) = .74$, $p < .001$; for Times 1 and 3, $r(62) = .61$, $p < .001$; and for Times 2 and 3, $r(64) = .61$, $p < .001$. Children's use of mental language was also consistent over time, although the correlations were of a smaller magnitude: for Times 1 and 2, $r(58) = .37$, $p < .01$; for Times 1 and 3, $r(62) = .46$, $p < .001$; and for Times 2 and 3, $r(62) = .47$, $p < .001$.

Performance on Theory-of-Mind Tasks

With few exceptions, at each time point, performance on the individual belief tasks (including the ambiguity task) was correlated with one another, as was performance on the different desire tasks (including the emotion–situation task), and the desire and belief tasks together. In addition, of the 12 possible

Table 3 Descriptive Statistics of Key Measures

	Time 1	Time 2	Time 3
	M (SD)	*M (SD)*	*M (SD)*
Theory of mind			
Transfer other	.10 (.30)	.36 (.48)	.76 (.43)
Transfer other justification	—	—	.49 (.50)
Contents self	—	—	.76 (.43)
Contents other	—	.25 (.43)	.43 (.50)
Contents other justification	—	—	.31 (.46)
Ambiguity	—	—	.21 (.41)
Desire–emotion	.37 (.49)	.57 (.50)	—
Emotion–situation	.57 (.22)	—	—
Desire–action	—	.42 (.50)	.67 (.47)
Wicked desires	—	—	.70 (.46)
Theory-of-mind composite	1.10/3.00 (.82)	1.62/4.00 (1.25)	4.43/8.00 (2.31)
Language			
M (*SD*)	8.94/20.00 (4.23)	11.86/20.00 (4.19)	15.56/20.00 (3.01)
Range	0–18	1–18	4–20
Mother mental state utterances			
M (*SD*)	18.26 (13.30)	35.22 (21.78)	32.32 (20.88)
Range	1–61	3–120	4–86
Child mental state utterances			
M	2.10 (3.71)	4.32 (6.10)	3.94 (4.05)
Range	0–27	0–34	0–19
Maternal education			
M	4.07 (1.51)	4.07 (1.51)	4.07 (1.51)
Range	0–6	0–6	0–6

Note: The mean proportion of correct answers is shown for each theory-of-mind task.

Table 4 Frequency of Mother Utterances

Category	Time 1	Time 2	Time 3
Mental state utterances			
Desire	3.46 (3.46)	3.33 (4.70)	2.50 (2.34)
Emotion	.84 (1.49)	5.61 (4.97)	4.93 (4.67)
Modulations of assertion	3.32 (3.72)	6.67 (7.53)	5.56 (5.78)
Think and know	8.22 (7.27)	17.07 (11.09)	15.69 (11.59)
Other mental state	1.54 (1.84)	1.15 (1.63)	1.32 (1.49)
Non-mental state utterances			
Descriptions	51.35 (17.74)	57.35 (20.65)	53.09 (20.86)
Elaborations of a theme	4.75 (5.31)	8.95 (8.56)	8.10 (6.75)
Causal talk	.58 (1.68)	1.68 (1.91)	1.51 (2.12)
Factual talk	.94 (1.47)	2.26 (2.60)	2.87 (3.44)
Links to child's life	12.41 (9.80)	9.97 (10.23)	8.31 (7.79)
Don't know	.49 (1.20)	.81 (1.33)	.60 (.97)
Physical state	.81 (1.75)	2.47 (2.17)	1.51 (1.59)
Orienting utterances	10.00 (5.12)	11.07 (6.30)	6.35 (4.45)
Repetitions of other	10.54 (6.96)	9.32 (7.43)	7.90 (7.20)
Self-repetitions	3.79 (3.55)	2.11 (2.19)	1.45 (1.70)

Note: Values in parentheses are standard deviations.

Table 5 Partial Correlations between Mothers' Mental State Utterances at Time 1 or 2 and Performance on Later Theory-of-Mind Tasks at Time 2 or 3 After Partialing out Earlier (Time 1 or 2) Theory-of-Mind Performance

Category	Time 1–2 ($n = 49$)	Time 1–3 ($n = 41$)	Time 2–3 ($n = 55$)
Mental state utterances			
Desire	.15	.09	.23*
Emotion	.11	.09	.05
Modulations of assertion	.24*	.22	.29*
Think and know	.38**	.13	.26*
Other mental state	.24*	.09	.09
Non-mental state utterances			
Descriptions	.28*	.10	.31**
Elaborations of a theme	.19	.21	.32**
Causal talk	.36**	.24$^+$	.01
Factual talk	.28*	−.14	.19
Links to child's life	.26*	.13	.23*
Don't know	.21	.07	.06
Physical state	.13	.22	.06
Orienting utterances	.01	−.08	.03
Repetitions of other	.17	−.07	.16
Self-repetitions	−.07	−.34*	−.04

* $p < .05$; ** $p < .01$; $^+p < .07$ (all significance tests are one-tailed).

correlations between belief and desire composites at different time points (e.g., Time 1 belief and Time 3 belief, Time 1 desire and Time 2 belief, and so forth), 11 were significant. Because of the high interrelatedness of the belief and desire tasks in this study and in other studies (e.g., Hughes & Dunn, 1998), performance was collapsed to form a theory-of-mind composite measure.

Relation between Mother Utterances and Theory of Mind

Table 5 lists the correlations between different mother utterances and the theory-of-mind composite. To examine whether mother utterances had any impact on subsequent theory of mind, correlations were computed between mother utterances at an earlier time point and theory-of-mind understanding at a later time point. In addition, theory-of-mind understanding at Time 1 was partialed out when examining correlations between Time 1 mother utterances and Time 2 or 3 theory-of-mind understanding. Time 2 theory-of-mind understanding was partialed out when examining correlations between Time 2 mother utterances and Time 3 theory-of-mind understanding.

There are several notable features of the data. Modulations of assertion and think and know terms correlated with later theory of mind at two of three sets of time points, desire and other mental state terms correlated with later theory of mind at one of three sets of time points, and emotion did not correlate with later theory of mind at any time point. In general, there was no one type of mental state utterance that exclusively correlated with theory-of-mind understanding. Among the non-mental state terms, descriptions, links to child's life, and causal talk (including one marginal correlation) correlated with later theory of mind at two of three sets of time points. Elaborations of a theme and factual talk correlated at one set of time points.

A second feature of the data was that there was only one significant negative correlation: self-repetitions. This makes perfect sense in that a mother would have to repeat herself if a child's overall ability was depressed or the child was inattentive and hence wouldn't learn much. The vast majority of correlations were positive and in the expected direction. Indeed, 13 of 14 significant correlations were positive, binomial test: $k = 13$, $n = 14$, $p < .001$. Thus, chance did not seem to play a role in determining the pattern of correlations.

Third, the likelihood of an utterance correlating with theory-of-mind understanding was not related to the frequency of an utterance. The frequency data in Table 4 and the correlations reported in Table 5 were used to compute correlations between (1) the mean frequency of an utterance, and (2) whether the utterance correlated significantly with later theory of mind. This was accomplished by using each of the

mental state utterances and the first five non-mental state utterances (descriptions, elaborations of a theme, causal talk, factual talk, and links to child's life). If utterances tended to be significant only when they were very frequent then a high positive correlation should have been obtained. The correlation, however, was only moderate and nonsignificant, $r(30) = .17$, *ns.*

Mental State versus Non-mental State Utterances

Examined next was whether mental state utterances and non-mental state utterances had effects on theory of mind independent of each other. Recall that a person's statement could be coded in multiple ways (e.g., as causal and mental state), so it is possible that non-mental state utterances correlated with theory of mind through shared variance with mental state utterances (or vice versa). A mental state composite variable (including all mental state utterances shown in Table 2) and a non-mental state composite variable (including descriptions, elaborations of a theme, causal talk, factual talk and links to child's life) were created. Table 6 includes the results of these analyses. In all three cases, early mental state usage facilitated later theory-of-mind understanding independent of early non-mental state usage. In contrast, early non-mental state usage never predicted unique variance. These data are consistent with the idea that only mother mental state utterances have a unique causal role in facilitating theory-of-mind understanding.

Within-Time Point Correlations

There were significant correlations within each time point between theory of mind and mother mental state utterances: Time 1, $r(49) = .50$, $p < .001$; Time 2, $r(63) = .38$, $p < .01$; and Time 3, $r(62) = .35$, $p < .01$; and between child mental state utterances and mother mental state utterances: Time 1, $r(67) = .47$, $p < .001$; Time 2, $r(65) = .60$, $p < .001$; and Time 3, $r(69) = .63$, $p < .001$. These correlations are consistent with the idea that mothers adjust their utterances on a minute time scale according to their children's level of skill. Another question (examined below) was whether children's language and theory of mind had any long-term impact on mother mental state utterances.

Relation between Mother Utterances and Children's Age, Theory of Mind, Language Ability, and Mental State Utterances

Children's age was not correlated with mothers' use of mental state utterances at Time 1, $r(49) = -.03$, *ns*: Time 2, $r(61) = .01$, *ns*, or Time 3, $r(60) = -.16$, *ns*. These findings helped to justify the analysis of group effects. Table 7 presents correlations between key variables and later theory-of-mind ability. The "mother mental state utterances" variable represents the composite of all mental state utterances. All variables correlated with later theory-of-mind understanding at a minimum of two sets of time points. Some of the correlations—particularly those with language, theory of mind, and mother mental state utterances—were notably high. The correlations in Table 7 do not specify the causal direction between variables, but highlight the necessity of partialing out early child mental state utterances, early language ability, and early theory of mind to determine whether mother mental state utterances have a unique causal role. Recall that this is something previous researchers did not do.

Assessing the Causal Role of Mother Mental State Utterances

Table 8 provides data on whether mother mental state utterances accounted for unique variance in theory-of-mind ability. One feature of the data was that mother mental state utterances had a large degree of variance at all time points, whereas theory-of-mind performance had a smaller amount of variance at all time points. Greater variance in the utterance variable might lead to the misleading conclusion that mother utterances had a unique causal role, whereas theory of mind did not. For this reason mother utterances were broken into quartiles when early mother mental state utterances and later theory-of-mind ability were examined. This had the added benefit of eliminating the influence of widely differing values on the corre-

Table 6 Partial Correlations between Early Mother Utterances and Later Theory-of-Mind Understanding

	Time 1–2	Time 1–3	Time 2–3
Early mental state utterances and later theory-of-mind understanding[a]	.50*** ($n = 63$)	.39** ($n = 54$)	.30* ($n = 56$)
Early non-mental state utterances and later theory-of-mind understanding[b]	−.19 ($n = 63$)	−.06 ($n = 54$)	.01 ($n = 56$)

[a] Partialing out early non-mental state utterances.
[b] Partialing out early mental state utterances.
* $p < .05$; ** $p < .01$; *** $p < .001$ (all significance tests are one-tailed).

Table 7 Correlates of Later Theory-of-Mind Understanding

	Time 1–2 (n = 48)	Time 1–3 (n = 40)	Time 2–3 (n = 52)
Earlier age	.31**	.24	.32*
Earlier mother mental state utterances	.60***	.41**	.50***
Earlier child mental state utterances	.45***	.32*	.09
Earlier language ability (CELF)	.60***	.59***	.57***
Earlier theory of mind	.63***	.54***	.66***
Maternal education (socioeconomic status)	.20	.38**	.34*

Note: CELF = Clinical Evaluation of Language Fundamentals.
*$p < .05$; **$p < .01$; *** $p < .001$ (all significance tests are one-tailed).

lations, which could affect the statistics when raw numbers are used. As a conservative measure, mother utterances were not broken into quartiles when correlations between early theory of mind and later mother mental state utterances were examined, so that there was more variance for theory of mind to account for. Thus, variance considerations meant that it was more likely that significant correlations between early theory of mind and later mother mental state utterances were to be found than the reverse.

Row 1 of Table 8 includes the correlations between early (e.g., Time 1) mother mental state utterances (quartiles) and later (e.g., Time 2) theory-of-mind understanding after partialing out early (e.g., Time 1) theory-of-mind understanding. Correlations at all three sets of time points were significant. Row 2 presents the correlations between early theory-of-mind understanding and later mother mental state utterances (raw numbers) after partialing out earlier mother mental state utterances (raw numbers). In this case none of the correlations were significant. The unidirectional nature of these correlations is consistent with mother utterances playing a causal role. Mother mental state utterances facilitated later theory of mind, but theory of mind did not shape later mother utterances.

Note that, above, control question failures were treated as test question failures on the grounds that a correct answer to a test question may have been misleading if a child failed a control question. An alternative strategy is to analyze test question performance irrespective of control question performance on the grounds that the child could have been distracted when asked control questions but their test question answer indicated genuine understanding. In line with this view, mother mental state utterances (quartiles) were examined as a predictor of later theory-of-mind test question performance (irrespective of control question performance), after partialing out early theory-of-mind test question performance. The pattern of results was identical. Time 1 mother mental state utterances were a significant predictor of Time 2 theory of mind, (partial correlation) $pr(48) = .37$, $p < .01$, and Time 3 theory of mind, $pr(44) = .40$, $p < .01$. Likewise, Time 2 mother mental state utterances were a significant predictor of Time 3 theory of mind, $pr(53) = .38$, $p < .01$.

Next, only the children at Time 1 who were actually given the false-belief task were included (recall that children under 3 years of age were not given the false-belief task at Time 1 and were assumed to have failed). These analyses resulted in an identical pattern of results. The correlation between Time 1 mother utterances and Time 2 theory of mind after partialing out Time 1 theory of mind was $pr(33) = .34$, $p < .05$. The correlation between Time 1 mother utterances and Time 3 theory of mind after partialing out Time 1 theory of mind was $pr(33) = .36$, $p < .05$. The correlation between Time 1 theory of mind and Time 2

Table 8 Partial Correlations between Mother Mental State Utterances, Child Mental State Utterances, Age, Language Ability, and Theory of Mind

	Time 1–2	Time 1–3	Time 2–3
Early mother mental state utterances (quartiles) and later theory-of-mind understanding[a]	.31** (n = 51)	.31** (n = 47)	.35** (n = 56)
Early theory-of-mind understanding and later mother mental state utterances[b]	−.06 (n = 51)	−.09 (n = 47)	−.02 (n = 56)
Early mother mental state utterances and later child mental state utterances[c]	.41** (n = 58)	.51*** (n = 62)	.39** (n = 63)
Early child mental state utterances and later mother mental state utterances[b]	−.12 (n = 58)	.12 (n = 62)	.09 (n = 63)

[a] Partialing out early theory-of-mind understanding.
[b] Partialing out early mother mental state utterances.
[c] Partialing out early child mental state utterances.
** $p < .01$; *** $p < .001$ (all significance tests are one-tailed).

mother mental state utterances after partialing out Time 1 mother mental state utterances was *pr*(33) = –.03, *ns.* The correlation between Time 1 theory of mind and Time 3 mother utterances after partialing out Time 1 mother utterances was *pr*(33) = –.18, *ns.* In an additional effort to standardize the variances, the theory-of-mind and mother mental state utterance values were transformed into *z* scores. Mother mental state utterances were a significant predictor of later theory of mind at all three sets of time points, *pr range* = .29–.34, all *p*s < .05, one-tailed. In contrast, theory of mind was never a significant predictor of later mother mental state utterances, *pr range* = –.02 to –.09.

The results also showed that at all three sets of time points, early mother utterances were linked to later child utterances, whereas early child utterances never predicted later mother utterances (Rows 3 and 4 of Table 8). These results are also consistent with a unique role for mother utterances.

Comparing Mother Utterances to Language Ability

Table 9 presents the results of linear regressions in which mother mental state utterances were compared directly to language ability to determine which accounted for more variance in theory-of-mind ability. Mothers' mental state utterances were predictors of later theory of mind after partialing out early theory of mind, early child mental state utterances, early language, maternal education (SES), and later age (because children differed in age at the later time point). Analogous analyses were carried out for language, after partialing out early mother mental state utterances. Having accounted for all such potentially confounding variables, there was still clear evidence that both mother utterances and language played a unique causal role in facilitating theory of mind. Indeed, mothers' mental state utterances correlated with subsequent theory of mind at all three sets of time points, and the average partial correlation was *pr* = .36 with theory of mind. Language correlated at two of three sets of time points, and the average partial correlation with theory of mind was *pr* = .29. Together, all variables accounted for between about half and two thirds of the variance in subsequent theory-of-mind performance. In sum, the results of the present study replicated the results of previous research in finding that language ability was highly related to theory-of-mind understanding, and added the new finding that mother mental state utterances explained a comparable or even greater amount of unique variance.

Think and Know Utterances versus Desire Utterances

Recall that Bartsch and Wellman (1995) found that parents' early desire talk was a stronger correlate of children's later talk about thoughts and knowledge

Table 9 Summary of Regression Statistics Predicting Theory-of-Mind Understanding

Variable	β	*t*	R^2 and ΔR^2	*pr*
Time 1 to Time 2 (n = 51)				
Step 1: Time 1 theory of mind, child MS utterances, mother MS utterances, maternal education, and Time 2 age			.49	
Step 2: Time 1 language	.18	1.31	.02	.19
Step 1: Time 1 theory of mind, child MS utterances, language, Time 2 age, and maternal education			.46	
Step 2: Time 1 mother MS utterances	.28	2.02*	.05	.29
Time 1 to Time 3 (n = 47)				
Step 1: Time 1 theory of mind, child MS utterances, mother MS utterances, maternal education, and Time 3 age			.44	
Step 2: Time 1 language	.39	2.86**	.10	.41
Step 1: Time 1 theory of mind, child MS utterances, language, Time 3 age, and maternal education			.48	
Step 2: Time 1 mother MS utterances	.31	2.21*	.06	.33
Time 2 to Time 3 (n = 54)				
Step 1: Time 2 theory of mind, child MS utterances, mother MS utterances, Time 3 age, and maternal education			.59	
Step 2: Time 2 language	.29	2.83**	.06	.38
Step 1: Time 2 theory of mind, child MS utterances, language, Time 3 age, and maternal education			.54	
Step 2: Time 2 mother MS utterances	.45	3.84***	.11	.49

Note: β = standardized regression coefficient; R^2 = proportion of variance explained by a variable; ΔR^2 = change in proportion of variance (i.e., additional variance explained by a variable); *pr* = partial correlation; MS = mental state.
* $p < .05$; ** $p < .01$; *** $p < .001$.

than was parents' early talk about thoughts and knowledge. These authors, however, did not use a standardized measure of children's language to determine if language mediated the relation. To explore this issue, the present study examined the earliest measure of mothers' talk (Time 1) and the next subsequent measure of children's talk (Time 2 talk about thoughts and knowledge), after partialing out Time 1 children's talk about thoughts and knowledge and general language ability at Time 1. The mean number of genuine desire utterances at Time 1 was 3.28 (SD = 3.28), and the mean number of genuine think and know utterances was 8.22 (SD = 7.27). The correlation between mothers' Time 1 genuine talk about desires and thoughts and knowledge was $r(53) = .46$, $p < .001$. The correlation between mothers' genuine desire utterances at Time 1 and children's genuine think and know utterances at Time 2 was $r(53) = .45$, $p < .001$, whereas the correlation between mothers' genuine think and know utterances at Time 1 and children's genuine think and know utterances at Time 2 was $r(53) = .24$, $p < .001$. The difference between these two correlation coefficients was significant, $t(53) = 1.65$, $p < .05$, one-tailed.

Next we examined whether children's genuine talk about desires preceded their genuine talk about thoughts and knowledge, by analyzing the data of the 56 children who were given the picture task at all three time points. Figure 1 includes information about the number of children and mothers who, at least once, provided genuine desire or think and know utterances. Over all three time points virtually all mothers provided at least one genuine think or know utterance, and most also provided at least one genuine desire utterance. In contrast, there was a steady increase in the number of children who provided genuine think and know utterances over the three time points. There were 16 children who provided at least one genuine think or know utterance at Time 2 but not at Time 1, and only 6 children who did the opposite, McNemar's $\chi^2(1, N = 22) = 4.55$, $p < .05$. In addition, 21 children provided at least one genuine think or know utterance at Time 3 but not at Time 2, and only 6 did the opposite, McNemar's $\chi^2(1, N = 27) = 8.33$, $p < .01$. In contrast, there was a slight decrease in the number of children who provided genuine desire utterances over the three time points. This drop was significant between Times 1 and 3: 15 children provided at least one genuine desire utterance at Time 1 but not at Time 3, and only 6 did the opposite, McNemar's $\chi^2(1, N = 21) = 3.86$, $p < .05$.

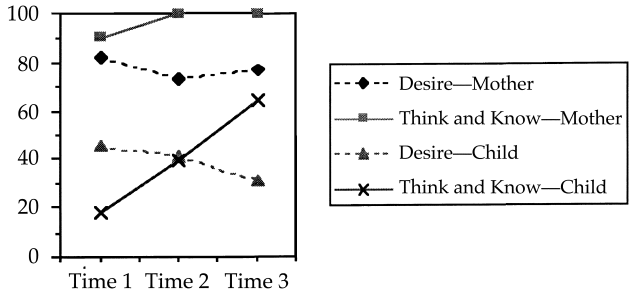

Figure 1 Percentage of mothers and children who provided at least one genuine desire or think and know response.

Furthermore, children were significantly more likely to provide genuine desire utterances at Time 1 than think or know utterances, but the pattern was reversed at Time 3. There were 23 children who provided at least one genuine desire utterance at Time 1 but no genuine think and know utterances, and only 4 children who did the reverse, McNemar's $\chi^2(1, N = 27) = 13.37$, $p < .001$. In addition, 29 children at Time 3 provided at least one genuine think and know utterance but no genuine desire utterances, and only 6 children did the reverse, McNemar's $\chi^2(1, N = 35) = 15.11$, $p < .001$. Like mothers, then, children eventually talked more about thoughts and knowledge than they did about desires. In addition, at each of the three time points, the number of mothers who provided genuine desire and think and know utterances far exceeded the number of children who did so. There was clear evidence, therefore, that children's lack of talk about thoughts and knowledge was not due to a lack of parental input about thoughts and knowledge.

GENERAL DISCUSSION

Previous research has not allowed strong inferences about whether mothers' utterances play a causal role in facilitating theory of mind, a point also made by Dunn and colleagues (Dunn, Brown, & Beardsall, 1991; Dunn, Brown, Slomkowski, et al., 1991). For instance, it was possible that mothers' utterances facilitated children's general language ability, which then enabled better theory-of-mind performance. It was also not possible to tell whether increases in mothers' mental state language were simply due to their picking up on an already-advanced theory of mind in children (manifest in the children's own mental state talk and theory-of-mind task performance); nor was it possible to tell whether it was mental state utterances per se that facilitated theory of mind, or some other kind of intervention that mothers used.

In the present study, mothers were asked to talk about a series of pictures with their children. Mothers' and children's language was examined for its relation

with theory-of-mind understanding. Mothers' mental state utterances were found to be correlated with children's mental state utterances and theory of mind within time points. This is consistent with the notion that mothers direct their interventions at a level appropriate to that of the child. The results of the present study, however, were clear in indicating a causal role for mothers' mental state utterances in facilitating children's subsequent theory of mind at all three sets of time points. These relations held even when many potentially mediating variables were accounted for, including the children's language ability, their initial social understanding (as manifest in their initial theory of mind and mental state utterances), their age, and the mothers' educational background. Further, the measure of theory of mind at early time points made it possible to determine that there was a unidirectional relation between mothers' mental state utterances at early time points and children's subsequent theory of mind. With these safeguards in place we can be more confident that mothers' mental state language plays a genuine causal role in facilitating theory of mind. This is particularly true given the great consistency across the three sets of time points in this study. Of course, it is necessary to include the proviso that the mothers in the present study were from middle- and upper middle-class areas. It is possible that mothers of working-class children do not use mental state utterances to the same extent, so that there are fewer benefits to be had for these children.

Another novel finding was the comparison between mothers' mental state utterances and language as predictors of theory of mind. Previous research has established that general language ability is one of the best correlates of theory of mind. Nevertheless, when mothers' mental state utterances were directly compared to language ability, mothers' mental state utterances more consistently accounted for unique variance in theory of mind. The language task used in the present study tapped semantic understanding. Although there have been claims that false belief is related primarily to syntactic abilities (Astington & Jenkins, 1999; de Villiers & de Villiers, 2000), there are many studies in which correlations have been found between theory of mind and measures of semantic understanding, such as the Peabody Picture Vocabulary Test or the British Picture Vocabulary Test (e.g., Happé, 1995; Hughes, 1998; Hughes and Dunn, 1997; Taylor & Carlson, 1997). The present study's finding of consistently high correlations between semantics and theory of mind is consistent with these past studies.

The present study went beyond previous research in making a comprehensive examination of mothers' non-mental state interventions, on the grounds that it was these, rather than mental state utterances, that facilitated theory of mind. Despite correlating with subsequent theory-of-mind ability, however, non-mental state utterances were not uniquely related to later belief understanding independent of mental state utterances. These findings have implications for previous studies. As in the study by Welch-Ross (1997), the present study found that links to the child's own life were related to theory of mind. The data, however, indicate that links had no special significance over and above mental state utterances. Similarly, as in previous research (Dunn, Slomkowski, et al., 1991), this study found that mothers' causal language correlated with subsequent theory of mind; however, causal language had no unique relation with theory of mind over and above that of mental state utterances. Thus, it is likely that it is only causal talk about mental states that is beneficial to theory of mind (Dunn, Brown, & Beardsall, 1991).

We deliberately used two different sets of pictures in the present study to enable more general claims. One set encouraged relatively few mental state utterances, and one encouraged many such utterances. The great consistency in the data at each set of time points demonstrates that the findings were not a function of a particular set of stimuli.

Several types of utterances were related to theory-of-mind understanding over the three sets of time points. The significant correlates included modulations of assertion at two sets of time points, which were thought particularly likely to be correlates of theory of mind for both theoretical reasons (they make specific reference to uncertainty) and empirical reasons (they have been found to correlate with belief understanding). Think and know terms also correlated with theory of mind at two sets of time points, as did desire terms and other mental state utterances at one set of time points. Nevertheless, the most consistent correlate of theory of mind was not any one category of utterance, but the composite variable of mothers' mental state utterances. This finding leads us to believe that it is general talk about mental states that helps children learn about the mind, rather than one specific type of utterance.

The present study's data also provided support for several aspects of Bartsch and Wellman's (1995) study of children's developing mental state language. The strength of their study was in the wealth of data obtained for individual children, but the limitations were in the small number of children examined, the absence of theory-of-mind task measures, and the absence of a standardized measure of children's language. The present study did not obtain such a rich catalogue

of children's language but did include a larger number of children, theory-of-mind measures, and measures of general language abilities. It is reassuring that despite these differences, consistent data were obtained. As in the Bartsch and Wellman study, the present study found that children talk about desires before they talk about thoughts and knowledge, and that this is true despite a wealth of early parental talk about thoughts and knowledge. Indeed, at Time 1 the number of utterances about thoughts and knowledge by mothers was more than double their talk about desires.

Further, consistent with the findings by Bartsch and Wellman (1995), the findings of the present study showed that mothers' talk about desires was more highly correlated with children's subsequent talk about thoughts and knowledge, than was mothers' talk about thoughts and knowledge. The measure of general language ability also allowed us to demonstrate that this effect was not simply a function of children's general language development. As Bartsch and Wellman suggest, children seem to learn about beliefs through mothers' discussion of desires. This finding is reminiscent of claims that training in desire facilitates belief understanding (Slaughter & Gopnik, 1996), and that mothers who discuss the feelings of others have children with enhanced belief understanding (Ruffman et al., 1999). Nevertheless, we must qualify such claims when theory-of-mind tasks themselves, rather than just children's language, are considered. Mothers' talk about thoughts and knowledge was associated with better theory-of-mind performance at two of three sets of time points, whereas mothers' talk about desires was associated with better theory of mind at only one set of time points. This leads us to believe that mothers' desire talk might provide some initial impetus for belief understanding by facilitating children's talk about thoughts and knowledge, but that mothers' general talk about all kinds of mental states may ultimately be the most important facilitator of theory-of-mind insights.

Further, it is possible that at some time earlier than examined by the present study or Bartsch and Wellman's (1995) study, mothers may talk more about desires with their children than they did about beliefs. Once their children begin to use desire terms mothers may then begin to talk more about beliefs, in an attempt to scaffold these insights. At present there are insufficient data to evaluate this idea, primarily because analyses of language have not begun early enough in a child's life and continued for long enough thereafter. Nevertheless, the data are roughly consistent. For instance, Bartsch and Wellman collected data for three children's parents when the children were 25 to 28 months, and then 29 to 32 months. At the earlier age, parents talked more about desires: a proportion of .49 for desires versus .30 for knowledge and thoughts. At the older age, desire talk and talk about thoughts and knowledge was about even: .39 for desires and .37 for knowledge and thoughts. Likewise, in the present study's data, mothers' talk about desires and thoughts and knowledge were comparable at Time 1, but by Times 2 and 3 there was more talk about thoughts and knowledge. A task for future research is to longitudinally examine parent language in children beginning at a younger age, to consolidate these ideas.

A final issue concerns how mothers' language facilitates children's theory of mind. For instance, does it provide them with theoretical knowledge (the theory theory), or help them take the perspective of others by using themselves as an analogy (the simulation theory)? If mothers primarily provide fact-based knowledge about other peoples' mental states, this might facilitate theoretical knowledge. If mothers frequently make links between the child's mental states and those of others, this might stimulate some kind of simulation processes (Ruffman et al., 1999). The present data do not address this issue, making it a question for future research.

In sum, the data of the present study are consistent with the general picture that mothers' language is related to both children's subsequent language and their theory of mind, and add to a growing body of research that indicates that parents as well as siblings are important for children's social understanding.

ACKNOWLEDGMENTS

The authors thank the Economic and Social Research Council for financial support (award #R000237071), and Claire Hughes for her very helpful advice.

ADDRESSES AND AFFILIATIONS

Corresponding author: Ted Ruffman, Experimental Psychology, University of Sussex, Brighton, East Sussex, BN1 9QG, United Kingdom; Lance Slade and Elena Crowe are also at the University of Sussex.

APPENDIX A

THE CLINICAL EVALUATION OF LANGUAGE FUNDAMENTALS–PRESCHOOL LINGUISTIC CONCEPTS SUBTEST

Item 1: Point to one of the bears. [shown four bears of different sizes, one of them is climbing]

Item 4: Point to a dog, but not the one that is eating. [shown three dogs, only one of who is eating]

Item 8: Point to an elephant next to a giraffe. [shown two elephants and three giraffes, only one elephant is next to a giraffe]
Item 12: Point to the turtle before you point to a fish. [shown four fish and one turtle]
Item 16: Point to the last bird in line. [shown four birds in line]
Item 20: Point to the giraffe after you point to an elephant and a monkey. [shown two elephants, a tiger, a giraffe, and two monkeys]

APPENDIX B

ITEMS OF THE EMOTION-SITUATIONS TASK

1. This is Sarah. This is Sarah's mummy. Sarah's mummy gives Sarah a great big present. (happy)
2. This is Johnny. Johnny wants to ride his truck but he can't. It's gone. Someone has taken it. (sad, angry)
3. Look, Sarah's mum is taking Sarah to the swings and slide. (happy)
4. Sarah's favorite toy is broken. (sad, angry)
5. There's a nasty dog with big teeth chasing Johnny. (scared)
6. Sarah has just built a big tower. She thinks it's really good. Johnny doesn't like it and pushes it over. (sad, angry)
7. Look, Johnny is in his paddling pool. (happy)
8. Johnny has taken Sarah's favorite toy and won't share it. (sad, angry)
9. Johnny can't find his football and he really wants to play with it. (sad, angry)
10. There's a tiger chasing Sarah. (scared)
11. This is Johnny. This is Johnny's sister. Johnny's sister gives Johnny an ice cream. (happy)
12. Johnny wants to play with Sarah and her friend but they won't let him. (sad, angry)
13. Johnny can't find his mummy at the swings and slide. (sad, scared)
14. Mummy dog can't find her puppy. (sad, scared)

Note: Correct answers (given by adults) are shown in parentheses.

REFERENCES

Astington, J. W., & Jenkins, J. M. (1995). Theory of mind development and social understanding. *Cognition and Emotion, 9*, 151–165.

Astington, J. W., & Jenkins, J. M. (1999). A longitudinal study of the relation between language and theory-of-mind development. *Developmental Psychology, 35*, 1311–1320.

Bartsch, K., & Wellman, H. M. (1995). *Children talk about the mind*. Oxford, U.K.: Oxford University Press.

Brown, J. R., Donelan-McCall, N., & Dunn, J. (1996). Why talk about mental states? The significance of children's conversations with friends, siblings, and mothers. *Child Development, 67*, 836–849.

Charman, T., & Shmueli-Goetz, Y. (1998). The relationship between theory of mind, language ability and narrative discourse: An experimental study. *Cahiers de Psychologie Cognitive, 17*, 245–271.

Churchland, P. M. (1988). *Matter and consciousness.* Cambridge, MA: MIT Press.

Clements, W. A., Rustin, C. L., & McCallum, S. (2000). Promoting the transition from implicit to explicit understanding: A training study of false belief. *Developmental Science, 3*, 81–92.

Cutting, A. L., & Dunn, J. (1999). Theory of mind, emotion understanding, language, and family background: Individual differences and interventions. *Child Development, 70*, 853–865.

Denham, S. A. (1986). Social cognition, prosocial behavior, and emotion in preschoolers: Contextual validation. *Child Development, 57*, 194–201.

de Villiers, J. G., & de Villiers, P. A. (2000). Linguistic determinism and the understanding of false beliefs. In P. Mitchell & K. Riggs (Eds.), *Children's reasoning and the mind* (pp.191–228). Hove, U.K.: Psychology Press.

Dunn, J., Brown, J., & Beardsall, L. (1991). Family talk about feeling states and children's later understanding of others' emotions. *Developmental Psychology, 27*, 448–455.

Dunn, J., Brown, J., Slomkowski, C., Tesla, C., & Youngblade, L. (1991). Young children's understanding of other people's feelings and beliefs: Individual differences and their antecedents. *Child Development, 62*, 1352–1366.

Gopnik, A. (1993). How we know our minds: The illusion of first-person knowledge of intentionality. *Behavioral and Brain Sciences, 16*, 1–14.

Gordon, R. M. (1992). The simulation theory: Objections and misconceptions. *Mind & Language, 7*, 11–34.

Happé, F. (1995). The role of age and verbal ability in the theory of mind task performance of subjects with autism. *Child Development, 66*, 843–855.

Harris, P. L. (1992). From simulation to folk psychology: The case for development. *Mind & Language, 7*, 120–144.

Hughes, C. (1998). Executive function in preschoolers: Links with theory of mind and verbal ability. *British Journal of Developmental Psychology, 16*, 233–253.

Hughes, C., & Dunn, J. (1997). "Pretend you didn't know": Preschoolers talk about mental states in pretend play. *Cognitive Development, 12*, 477–499.

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.

Lewis, C., Freeman, N., Kyriakidou, C., Maridaki-Kassotaki, K., & Berridge, D. (1996). Social influences on false belief access: Sibling influences or general apprenticeship? *Child Development, 67*, 2930–2947.

Lewis, C., & Mitchell, P. M. (1994). *Children's early understanding of mind.* Hove, U.K.: Erlbaum.

Meins, E., & Fernyhough, C. (1998). Linguistic aquisitional style and mentalising development: The role of maternal mind-mindedness. *Cognitive Development, 14*, 363–380.

Meins, E., Fernyhough, C., Russell, J., & Clark-Carter, D. (1998). Security of attachment as a predictor of symbolic and mentalising abilities: A longitudinal study. *Social Development, 7*, 1–24.

Moore, C., Jarrold, C., Russell, J., Lumb, A., Sapp, F., & MacCallum, F. (1995). Conflicting desire and the child's theory of mind. *Cognitive Development, 10*, 467–482.

Moore, C., Pure, K., & Furrow, D. (1990). Children's understanding of the modal expression of speaker certainty and uncertainty and its relation to the development of a representational theory of mind. *Child Development, 61*, 722–730.

Perner, J. (1991). *Understanding the representational mind.* Cambridge, MA: MIT Press.

Perner, J., Leekam, S. R., & Wimmer, H. (1987). Three-year-olds' difficulty with false belief: The case for a conceptual deficit. *British Journal of Developmental Psychology, 5*, 125–137.

Perner, J., Ruffman, T., & Leekam, S. R. (1994). Theory of mind is contagious: You catch it from your sibs. *Child Development, 65*, 1228–1238.

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.

Shatz, M., Wellman, H. M., & Silber, S. (1983). The acquisition of mental verbs: A systematic investigation of the first reference to mental state. *Cognition, 14*, 301–321.

Slaughter, V., & Gopnik, A. (1996). Conceptual coherence in the child's theory of mind: Training children to understand belief. *Child Development, 67*, 2967–2988.

Taylor, M. (1988). Conceptual perspective taking: Children's ability to distinguish what they know from what they see. *Child Development, 59*, 703–718.

Taylor, M., & Carlson, S. M. (1997). The relation between individual differences in fantasy and theory of mind. *Child Development, 68*, 436–455.

Thorpe, K. (1996). Thorpe Interaction Measure (TIM). Unpublished manuscript, University of Bristol, Bristol, U.K.

Welch-Ross, M. K. (1997). Mother–child participation in conversation about the past: Relationships to preschoolers' theory of mind. *Developmental Psychology, 33*, 618–629.

Wellman, H. M. (1990). *The child's theory of mind.* Cambridge, MA: MIT Press.

Wellman, H. M., & Bartsch, K. (1988). Young children's reasoning about beliefs. *Cognition, 30*, 239–277.

Wellman, H. M., Harris, P. L., Banerjee, M., & Sinclair, A. (1995). Early understanding of emotion: Evidence from natural language. *Cognition and Emotion, 9*, 117–149.

Wellman, H. M., & Woolley, J. D. (1990). From simple desires to ordinary beliefs: The early development of everyday psychology. *Cognition, 35*, 245–275.

Wiig, E. H., Secord, W., & Semel, E. (1992). *Clinical Evaluation of Language Fundamentals–Preschool.* San Antonio, TX: The Psychological Corporation.

Wimmer, H. M., & Mayringer, H. (1998). False belief understanding in young children: Explanations do not develop before predictions. *International Journal of Behavioral Development, 22*, 403–422.

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.

Yuill, N., Perner, J., Pearson, A., Peerbhoy, D., & van den Ende, J. (1996). Children's changing understanding of wicked desires: From objective to subjective and moral. *British Journal of Developmental Psychology, 14*, 457–475.

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.