Pathways to Understanding Mind: Construct Validity and Predictive Validity of Maternal Mind-Mindedness

Elizabeth Meins, Charles Fernyhough, Rachel Wainwright, David Clark-Carter, Mani Das Gupta, Emma Fradley, and Michelle Tuckey

The construct validity of maternal mind-mindedness (MM) was investigated in the context of its relations with children’s later understanding of mind. MM measures were obtained from infant–mother (N = 52) interactions at 6 months, and from maternal interviews at 48 months. Children’s understanding of mind was assessed using theory of mind (ToM) tasks at 45 and 48 months, and a stream of consciousness (SoC) task at 55 months. One of the early MM measures—mothers’ appropriate mind-related comments—was a positive independent predictor of: (a) MM at 48 months, and (b) ToM and SoC performance at 45 to 55 months. Path analyses suggested direct links between mothers’ use of appropriate mind-related comments and children’s later understanding of mind.

There is now an impressive body of evidence for social and environmental influences on young children’s developing understanding of mind. In their landmark study, Dunn, Brown, Slomkowski, Tesla, and Youngblade (1991) found that children were more likely to pass age-appropriate theory of mind (ToM) tasks if their families had a tendency to discuss feelings and use causal mental state language, and if their mothers regularly attempted to control the behavior of older siblings. Subsequent studies have found that children are more successful on ToM tasks if they have more siblings (Jenkins & Astington, 1996; Perner, Ruffman, & Leekam, 1994) and greater contact with older kin (Lewis, Freeman, Kyriakidou, Maridaki-Kassotaki, & Berridge, 1996; Ruffman, Perner, Naito, Parkin, & Clements, 1998); if they were securely attached in infancy (Meins, Fernyhough, Russell, & Clark-Carter, 1998), if their mothers tend to use mental state language during a picture-describing task (Ruffman, Slade, & Crowe, 2002), and if children themselves commonly use mental state language in conversations with siblings and friends (Brown, Donelan-McCall, & Dunn, 1996).

It has recently been suggested that at least some of these social-environmental correlates of ToM development might be related to caregivers’ mind-mindedness (MM): the proclivity to treat one’s infant as an individual with a mind, rather than merely an entity with needs that must be satisfied (Meins, 1997). The construct of MM thus builds on a tradition of research that has attributed an important role in mental development to caregivers’ tendency to treat infants as intentional agents (Bruner, 1975; Fonagy, Steele, Steele, Moran, & Higgitt, 1991; Newson, 1979; Trevarthen, 1979). In their longitudinal study, Meins et al. (1998) reported that mothers of securely attached children were more likely than their insecure group counterparts to focus on their children’s mental attributes (rather than physical appearance or behavioral tendencies) when asked to describe them at age 3. This measure of maternal MM was also related to children’s performance on a belief-preference integration ToM task (Harris, Johnson, Hutton, Andrews, & Cooke, 1989) at age 5. On this evidence, Meins et al. (1998) suggested that the apparent link between ToM development and security of attachment might be explicable in terms of individual differences in maternal MM.

In a subsequent study, Meins, Fernyhough, Fradley, and Tuckey (2001) asked whether indivi-
dual differences in maternal MM could be identified in the earliest months of life. From observations of mother–infant free-play interactions at 6 months, they identified five classes of maternal behavior that could be seen as manifestations of MM: maternal responsiveness to change in infant’s direction of gaze, maternal responsiveness to infant’s object-directed action, imitation, encouragement of autonomy, and appropriate mind-related comments. One of these measures, appropriate mind-related comments, in turn predicted security of attachment as measured using the strange situation at 12 months. Furthermore, this measure of maternal mind-oriented language proved to be a better predictor of security than maternal sensitivity (Ainsworth, Bell, & Stayton, 1971). In interpreting their findings, Meins et al. (2001) speculated that continuity in maternal MM between infancy and early childhood might facilitate children’s acquisition of a representational ToM and thus account for security-related differences in mentalizing abilities (Meins et al., 1998). Some support for this view was presented by Meins and Fernyhough (1999), who considered mothers’ attribution of meaning to infants’ apparently meaningless utterances at 20 months as a measure of mothers’ proclivity to treat their infants as intentional agents. They found that the incidence of such attributions was related to children’s later performance on a belief-preference integration ToM task (Harris et al., 1989).

Stronger evidence for a link between early maternal MM and later ToM came from the longitudinal study of Meins et al. (2002). They reported that mothers’ appropriate comments on their infants’ mental states at 6 months predicted children’s performance on a battery of ToM tasks at 45 and 48 months. Moreover, these researchers demonstrated that previously identified social-environmental correlates of ToM (maternal education, number of siblings, exposure to general mental state language, and attachment security) were not independent predictors once early maternal MM had been taken into account. In addition, this study showed that the value of early MM as a predictor of subsequent ToM performance was not due to its indexing of some general quality of early infant–mother interaction, since maternal sensitivity (Ainsworth et al., 1971) at 6 months was not found to be an independent predictor of ToM. Nor was there any relation between later ToM performance and early exposure to mental state language that did not appropriately reflect the infant’s current state of mind.

Two important questions arise from these investigations into the construct of maternal MM. The first relates to Meins et al.’s (2001) original identification of five maternal behaviors that might in principle be taken as manifestations of maternal MM. Only one of these behaviors, appropriate mind-related comments, was found to relate to attachment security. For this reason, it was the MM measure employed in Meins et al.’s (2002) study on longitudinal predictors of ToM. What distinguishes this measure from the other four MM measures considered by Meins et al. (2001) is that it draws on mothers’ explicit representations of their infants’ mental states. This feature is also shared with the MM measure used by Meins et al. (1998) in their original study, namely, the proportion of mental state attributes used by mothers in describing their children. Although these two measures have been successful in predicting later ToM performance, it remains to be seen whether the construct of MM can be reduced to caregivers’ tendency to produce explicit mentalistic representations of their infants, or whether the other elements of MM identified by Meins et al. (2001) also have their part to play. For us to be sure about the importance of MM as a predictor of ToM development, it is necessary to determine the predictive validity of these other four measures of MM that do not involve explicit linguistic representations of infants’ mental states.

The second question relates to the temporal stability of MM. In discussing their findings, Meins et al. (2002) considered two possible developmental pathways through which MM might have an effect on ToM development. First, MM in infancy might have a direct influence on the emergence of very early forms of ToM understanding, which in turn relate to ToM performance in the preschool years. Second, MM in infancy might have its effect through being continuous with MM in the preschool years, when the acquisition of a representational ToM is in full flow. Another way of putting this question is to ask whether there is a sensitive period for exposure to appropriate mental state language and other forms of MM behavior. If so, are the crucial advances made in infancy, or does MM not have its effect until the preschool years?

One way of addressing these questions is to consider the validity of the construct of MM. This was the primary aim of the longitudinal study reported here. Two aspects of validity are important for our purposes: construct validity (convergent and divergent) and predictive validity. We sought to address the question of convergent validity by first investigating the interrelations between the indexes of very early MM identified by Meins et al. (2001). Following Meins et al. (2002), we also investigated...
the category of maternal comments that were judged not to be appropriate reflections of the infant’s current state of mind. Our intention was to determine whether such comments were also indicative of MM (through a negative relation) or were unrelated to other indexes of MM.

The convergent validity of MM was also investigated by looking at the relation between early and later measures of MM. Investigating the temporal continuity of MM allowed us to ask whether it is early or later MM that is important in facilitating the acquisition of a representational ToM. The issue of divergent validity was investigated by comparing the predictive power of these measures of MM with that of another important index of early infant—mother interaction, maternal sensitivity (Ainsworth et al., 1971). The predictive validity of MM was investigated with respect to performance on standard ToM tasks and on a task designed to assess children’s understanding of the stream of consciousness (Flavell, Green, & Flavell, 1993).

Use of the stream of consciousness (SoC) task allows one to go beyond the traditional ToM focus on false-belief prediction and explanation and instead to look at children’s understanding of ongoing mental activities and processes. In their original study, Flavell et al. (1993) sought to establish whether children would acknowledge the existence of a person’s mental activity when outward clues were absent. Using a nonverbal response procedure, in which children had to choose a face with a thought bubble depicting a mind containing “some thoughts or ideas” or a mind “empty of thoughts or ideas,” Flavell et al. investigated whether children would attribute thoughts to an experimenter who was sitting quietly and “waiting.” They found that preschool children rarely acknowledged the SoC under these circumstances, with only 10% to 15% of 3-year-olds and 35% to 50% of 4-year-olds choosing the “some thoughts or ideas” face when the experimenter was “waiting.” Indeed, not even all of the adult participants in this study acknowledged the SoC in the absence of external manifestations of mental activity. In contrast, 70% of 3-year-olds and 90% of 4-year-olds maintained that the experimenter was thinking while looking at posters on a wall, and 65% of both 3- and 4-year-olds attributed mental activity to her when she was involved in solving a puzzle.

Flavell and colleagues have also reported that preschool-age children have problems in introspecting on their own mental processes. Although 3-year-olds understand that thinking is a purely internal activity different from physical and perceptual activities (Flavell, Green, & Flavell, 1995), they will frequently deny that they have been thinking, despite the fact that they have been engaged in obvious thought-provoking activity (Flavell et al., 1995; Flavell, Green, Flavell, & Grossman, 1997). Taken together, these findings suggest that, at an age when children typically pass standard ToM tasks to assess their understanding of mental states, they appear to have considerable difficulties in understanding ongoing human mental activity. Nonetheless, there are good reasons for thinking that these abilities will be related to one another. For example, recognizing that people can hold a false belief, and have different mental states from one’s own, might prompt children to question why people hold different beliefs, thus marking a first step toward their introspecting on how thoughts come about. Conversely, an understanding that mental activity continues in the absence of external behavior may help children understand why people act in the absence of obvious clues. No study has attempted to replicate Flavell et al.’s (1993) findings nor to investigate the relation between children’s basic ToM performance and their SoC understanding. Addressing this gap in the literature was one of the aims of the study reported here.

In addition to our interest in the link between ToM and SoC understanding, the present study was motivated by a number of reasons for supposing that SoC understanding would relate to early maternal MM. Flavell et al. (1993) proposed that one potential reason for the lag in children’s awareness of the SoC was that other people do not remark on it. The research on MM, however, has shown that certain parents do comment on such mental processes, even in the first year of life. For example, whereas some mothers never commented on their 6-month-olds’ minds, others frequently talked about their infants’ desires and preferences, and they attributed thoughts, memories, and even epistemic states to their infants (Meins et al., 2001). Exposure to MM discourse would thus seem to give children plenty of opportunities for reflecting on their own and others’ ongoing ideation, which may well be manifested in a relation between early MM and later SoC understanding.

Our specific hypotheses were as follows: (a) that the measures of early MM identified by Meins et al. (2001) would be significantly correlated with each other and with an interview-based measure of MM at age 4, (b) that both early and later MM would relate to a composite measure of ToM understanding at 45 and 48 months, (c) that both early and later MM would predict SoC understanding at 55 months, (d) that both early and later MM would be better
predictors of ToM and SoC understanding than a measure of maternal sensitivity, and (e) that scores on a composite measure of ToM at 45 and 48 months would relate to SoC understanding at 55 months. Finally, possible developmental pathways linking these measures were investigated using path analyses.

Method
Overview

Children took part in a longitudinal study between ages 6 and 55 months. The study was designed to investigate links between early and later MM, children’s ToM performance, and their acknowledgement of the SoC. At the 6-month testing stage, videotaped infant–mother interactions were coded for maternal MM and maternal sensitivity. At 45 months, children were given the appearance–reality (Flavell, Flavell, & Green, 1983) and deceptive box (Hogrefe, Wimmer, & Perner, 1986) ToM tasks. At 48 months children were given the unexpected transfer ToM task (Wimmer & Perner, 1983), and a measure of MM was obtained from a maternal interview (Meins et al., 1998). Children’s receptive verbal ability was also assessed at this age. Finally, the SoC task (Flavell et al., 1993) was administered at 55 months.

The longitudinal nature of the study reported here, coupled with our interest in the long-term predictive validity of the various MM measures (and particularly whether certain variables have a mediational role in developmental pathways to ToM), means that some of the intercorrelations reported here overlap with previously published reports on our original sample (Meins et al., 2001; Meins et al., 2002). The intercorrelations between the 6-month variables, with the exception of those involving inappropriate mind-related comments, were reported in Meins et al. (2001), albeit on the larger original sample of 71 infant–mother pairs. Intercorrelations between ToM performance and certain 6-month measures (appropriate and inappropriate mind-related comments, maternal sensitivity, and maternal education) were reported in Meins et al. (2002), albeit on a larger subset of the original sample. These previously published findings are included here to create a complete picture of how these variables interact in the development of ToM and SoC understanding, and discussion of them is kept to a minimum.

Participants

Participants were 52 children (25 girls, 27 boys) and their mothers. Three of the children were mixed race, and the remainder were White. The participating families were from predominantly lower middle-class backgrounds and lived in the Midlands region of the United Kingdom. As part of an earlier longitudinal study of 71 infant–mother pairs (Meins et al., 2001), measures of maternal MM and maternal sensitivity were obtained at 6 months (M age = 25 weeks; range = 23–28 weeks), and 52 of the original sample were followed up at the 45-month (M age = 45.80 months; range = 45–47 months), 48-month (M age = 48.25 months; range = 48–53 months), and 55-month (M age = 55.00 months, range = 54–58 months) testing stages. The reduction in numbers between the 6-month and follow-up phases was caused by 11 families moving away from the area, 5 mothers being unable to continue because of time constraints, and 3 families being impossible to contact.

Maternal Education

Mothers’ level of education was included as an independent variable because it has been found to correlate positively with children’s ToM understanding (Cutting & Dunn, 1999; Meins & Fernyhough, 1999; Meins et al., 2002). Mothers were awarded one of the following scores for their highest educational level (North American equivalents to the British educational system are given in parentheses): 0, no examinations; 1, CSEs (equivalent to high school up to age 16 for less academic students); 2, GCSEs or O-Levels (high school up to age 16 for more academic students); 3, A-Levels (high school up to age 18); 4, further qualification, not to degree level (e.g., nursing); 5, undergraduate degree; and 6, postgraduate qualification. Of the 52 mothers, 4 scored 0, 6 scored 1, 14 scored 2, 3 scored 3, 18 scored 4, 4 scored 5, and 3 scored 6.

Assessment of Maternal MM at 6 Months

Infant–mother interactions were videotaped when the infants were 6 months of age. These interactions consisted of a 20-min free-play session at the university’s developmental research laboratory. Mothers were given no specific instructions on how to act but rather were invited to play with their infants as they would if they had a few spare minutes together at home. Three types of maternal behavior were coded: (a) mothers’ responses to changes in their infants’ line of gaze, (b) mothers’ responses to their infants’ object-directed action, and (c) mothers’ vocal comments. Depending on the length of the utterance, a comment was defined as a
discrete sound, single word, or sentence. These behaviors were then classified for MM according to Meins et al.’s (2001) system of five categories: maternal responsiveness to change in infant’s direction of gaze, maternal responsiveness to infant’s object-directed action, imitation, encouragement of autonomy, and mind-related comments.

For the first two of these categories, responsiveness was defined as a mother looking at the object on which the infant’s attention or action was focused, touching or picking up the object, or naming or describing it. Behaviors falling into these two categories were coded directly from the videotaped infant–mother interactions. Scores for both of these categories were proportional. For maternal responsiveness to change in infant’s direction of gaze, each mother received a score for the number of responses made to changes in her infant’s line of gaze as a proportion of the total number of gaze changes made by her infant. For maternal responsiveness to infant’s object-directed action, each mother received a score for the number of responses made to her infant’s object-directed actions as a proportion of the total number of object-directed actions performed by her infant.

The three remaining MM categories related to mothers’ vocal comments during the sessions. Behaviors falling into these categories were coded directly from the videotapes of the infant–mother interactions. Imitation consisted of all maternal comments that were precise imitations of vocalizations produced by the infant, and scores for this variable were computed as a proportion of the total number of infant vocalizations made. Encouragement of autonomy consisted of all maternal comments that encouraged infants to perform an action independently, such as rolling over, sitting up, reaching for a toy, and so on.

The final category of MM was mind-related comments. Following Meins et al. (2001), mind-related comments were defined as: (a) comments on the infant’s mental state, such as knowledge, thoughts, desires, and interests (e.g., “An octopus—you know what one of those is,” “You want the cow”); (b) comments on the infant’s mental processes (e.g., “Do you recognize this?” “Are you making a decision?”); (c) references to the infant’s level of emotional engagement (e.g., commenting that the infant was bored, self-conscious, or excited); (d) comments on the infant’s attempts to manipulate other people’s beliefs (e.g., “You’re joking,” “You’re just teasing me”); and (e) the mother “putting words into her infant’s mouth” so that her discourse took the form of a dialogue.

Because infants’ mental states and processes are not transparent in this context, mothers must infer their infants’ likely state of mind. It is therefore necessary to consider the possibility that some mind-related comments may be inappropriate interpretations of infants’ states of mind. Consequently, Meins et al. (2001) argued that it was necessary to obtain an independent measure of whether each mind-related comment was appropriate or inappropriate.

Each mind-related comment was coded dichotomously as appropriate or inappropriate by a researcher who was blind to all other measures and to the study’s hypotheses. These categories were coded from the videotaped interactions with the aid of transcripts of mothers’ mental state talk. Comments were coded as appropriate if any of the following criteria were met: (a) the independent coder agreed with the mother’s reading of her infant’s psychological state; (b) the comment linked the infant’s current activity with similar events in the past or future; or (c) the comment clarified how to proceed if there was a lull in the interaction.

Comments were coded as inappropriate if any of the following criteria were met: (a) the coder believed that the mother was misinterpreting her infant’s psychological state; (b) the comment referred to a past or future event that had no obvious relation to the infant’s current activity; (c) the mother asked what the infant wanted to do, or commented that the infant wanted or preferred a different object or activity, when the infant was already actively engaged in an activity or was showing a clear preference for a particular object; or (d) the referent of the mother’s comment was not clear. A second researcher who was blind to all other measures and the hypotheses coded a random fifth of the tapes for the appropriate–inappropriate distinction. Interrater agreement was $\kappa = 0.79$. Higher scores for appropriate mind-related comments were taken to indicate greater MM.

In addition to the five MM categories, there was a miscellaneous category for comments not deemed to be MM (including comments that named or described objects, gave positive feedback on the child’s behavior, directed the child’s attention, and standardized game routines or rhymes). Mothers’ comments were coded by a researcher who was blind to all other measures and to the study’s hypotheses, with a randomly chosen fifth of the videotapes being coded by a second blind researcher. Interrater agreement using the five MM categories and the miscellaneous category was $\kappa = 0.90$.

To control for maternal verbosity, scores in the encouragement of autonomy, appropriate mind-
related comments, and inappropriate mind-related comments categories were calculated as a proportion of the total number of maternal comments produced during the session. It could be argued that frequency rather than proportional scores might provide a truer measure of children's exposure to mental state language. However, we maintain that any interesting relations between the frequency of mothers' mind-related comments and other variables might be confounded by the likelihood that more verbose mothers will also produce more inappropriate comments, thus diluting any beneficial effect of exposure to MM language. Alternatively, it may be that infants benefit from exposure to any kind of maternal language, MM or otherwise. Without controlling for maternal verbosity in our measures of MM, it would be impossible to discount these alternatives. That said, the following analyses using proportional scores produced an identical pattern of relations among the MM indexes, ToM performance, and SoC understanding when they were repeated using frequency measures. Only the analyses using proportional MM scores are therefore reported.

Assessment of Maternal Sensitivity at 6 Months

The 20-min infant–mother play sessions were also used to obtain a measure of maternal sensitivity using Ainsworth et al.'s (1971) 9-point scale, which gives a global rating of mothers' sensitivity to their infants' cues. Higher scores on this scale are indicative of more sensitive mothering. The videotaped infant–mother interactions were coded by a trained researcher, and a randomly chosen fifth of the tapes was coded by a second researcher. Both researchers were blind to all other measures and to the study's hypotheses. Interrater agreement was $k = 0.75$, with exact agreement for 79% of the observations.

Assessment of Maternal MM at 48 Months

When their children were 48 months of age, mothers were given a version of Meins et al.'s (1998) MM interview (note that one of the mothers did not complete the interview, giving a sample size of $N = 51$). The interview was conducted in the home after the child had completed the other tasks that were conducted at this age. The interview consisted simply of mothers being asked one question: “Can you describe [child] for me?” If mothers sought guidance on the kind of answer that was required, they were told that there were no right or wrong answers to the question and that they could talk about any of their children's characteristics they liked. The mothers' answers to this question were audiotaped and later transcribed verbatim. These transcripts were then coded for maternal MM using Meins et al.'s (1998) scheme, whereby each attribute mentioned was placed into one of four exhaustive and exclusive categories.

1. Mental. Any reference to the child’s mental life, relating to will, mind, imagination, interest, intellect, or metacognition (e.g., “bright,” “imaginative,” “She knows what she wants,” “He's always aware of other people's feelings,” “She's got a mind of her own”). Any comment relating to desires, wishes, and emotions (but not merely in terms of describing the child’s behavioral likes and dislikes)—for example, “He’s interested in dinosaurs”—was classified as a mental characteristic, but “She loves playing” was classified as a behavioral characteristic.

2. Behavioral. Any reference to behavior, such as games, activities, and interactions with others on a behavioral level (e.g., “He interferes with what his brother and sister are doing,” “She’ll go and mix with anybody”). The following descriptions were also categorized as behavioral because a purely nonmentalistic interpretation was possible: bubbly, lively, well mannered, cheerful, fun, placid, vibrant, outgoing, naughty.

3. Physical. Any physical characteristic was included, such as the child’s age and descriptions relating to the child’s position in the family (e.g., “brown hair,” “beautiful,” “quite big for his age”).

4. General. Any comments relating to the child that did not fit into the preceding three categories (e.g., “typical lad,” “lovely,” “he’s a bit of a pleasure”) were considered general.

Each mother received a score for the use of mental characteristics, expressed as a proportion of the total number of attributes used to describe the child. Higher scores on the mental category are indicative of greater MM. Once again, analyses using frequency scores produced an identical pattern of results to those using proportional scores; therefore, only the latter are reported. The interviews were coded by a researcher who was blind to all other measures and the study's hypotheses, and a second blind researcher coded a randomly selected fifth of the interviews. Interrater agreement was $k = 0.87$. 
ToM Performance

Children were tested at home on three ToM tasks: the appearance–reality task (Flavell et al., 1983), the deceptive box task (Hogrefe et al., 1986), and the unexpected transfer task (Wimmer & Perner, 1983). Given the finding that appearance–reality and informational-access competence typically emerges slightly before false-belief prediction performance (e.g., Carlson & Moses, 2001), the first two of these tasks were administered at 45 months, and the third at 48 months.

The Appearance–Reality Task

Children were given four trials, each using an object whose appearance was deceptive (a sponge that looked like a football, a torch that looked like a fish, a frog pencil sharpener, and a cat-shaped salt shaker). The experimenter showed the object to the child, saying, “When you look at this with your eyes right now, what does it look like?” After the child answered, the experimenter demonstrated what the object really was. The child was then asked two test questions: (a) “What is this really and truly?” and (b) “When you look at it with your eyes right now, does it look like a [football] or does it look like a [sponge]?” The order of presentation of the four objects, and that of the “look” versus “really and truly” questions, was fully randomized and counterbalanced. Children received 1 point if they answered both the reality and appearance questions correctly for each of the four objects, giving a score out of 4 for their overall performance on the appearance–reality task.

The Deceptive Box Task

Children were shown a tube of candies and asked what they thought was inside. The tube was then opened to show the child that it contained, not candies, but a pencil. The lid was replaced and the child was then asked a memory control question: “Can you remember what’s inside here?” After passing this control question, the child was introduced to a toy animal and asked to predict what the animal would think was in the tube. Performance was scored dichotomously as pass–fail.

The Unexpected Transfer Task

Children received two versions of this task, each using two different toy animals and two differently colored boxes. In one story, participants were introduced to Charlie the Crocodile and told that his favorite food was chocolate. A chocolate was placed into one of the boxes, and the child was told that Charlie was hiding his chocolate while he went for a swim. Charlie was removed from the scene, and a puppet, Cheeky Monkey, was introduced to the child. The experimenter announced that Cheeky Monkey was going to play a trick on Charlie. Cheeky Monkey then took the chocolate out of the box in which it had been hidden and placed it in the other box, closing both lids. The child was told that Charlie was about to return from his swim and that he would be wanting his chocolate. Two control questions were asked: (a) “Where was the chocolate in the beginning?” (memory control) and (b) “Where is the chocolate now?” (reality control). If a child answered either of these questions incorrectly, the story was briefly recapped and the two control questions repeated, but the child was not explicitly corrected. When correct answers had been given to both control questions, the test question was presented: “Where will Charlie look for his chocolate?” The second version of this task involved a second animal hiding a toy in one of two different boxes. For each of the two versions, children received a score of 1 if they passed or 0 if they failed, giving an overall score between 0 and 2.

Overall ToM Performance

A composite ToM score was computed for each child. So that equal weighting was given to each of the three tasks, scores for the individual tasks were scaled to give a maximum score of 2 for each task. Scores for the appearance–reality task were therefore divided by 2, and scores for the deceptive box were multiplied by 2. The maximum possible score for overall ToM performance was 6. The mean composite ToM score was 3.01 (SD = 1.90; range = 0–6). Skewness was -0.05, with a standard error of 0.26, $z = 0.10, ns$, and kurtosis for the composite measure was 1.10, $z = 1.62, ns$, indicating that composite ToM scores were normally distributed. Regarding reliability of the composite measure, all interitem and item–total correlations were positive and significant ($rs$ between 0.33 and 0.82), except that between performance on the unexpected transfer and appearance–reality tasks, $r(50) = 0.16$. The internal reliability of the composite measure was satisfactory (Cronbach’s $\alpha = 0.65$).

Children’s Receptive Verbal Ability

Children’s receptive verbal ability was assessed at home at 48 months, using the British Picture
Children’s Understanding of the SoC

At 55 months, participants received a version of Flavell et al.’s (1993) SoC task, the aim of which is to establish whether children attribute thoughts to a person engaged in various activities. This age was chosen to correspond as closely as possible to the mean age (54 months) of the 4-year-olds in Flavell et al.’s original study. The task was administered in the university’s developmental laboratory and required children to make nonverbal responses to test questions by choosing one of two schematic faces: one depicting a person who was thinking (a face with a thought bubble containing three asterisks) and the other showing a person whose mind was empty of thoughts or ideas (a face with an empty thought bubble).

Following Flavell et al. (1993), the testing session began with a procedure to familiarize children with the schematic faces. This involved the two experimenters (E1 and E2) describing the type of thought process that each of the schematic faces was intended to represent. The child was then asked to identify the picture that signified the mind “empty of thoughts or ideas” or the mind having “some thoughts or ideas.” If the child was unsure, the relevant information and questions were briefly repeated. None of the children had any problems in correctly identifying the different thought processes that the two faces were intended to convey. Children then participated in a warm-up procedure to ensure that they could use the schematic faces to label the mental processes of a real person. No participants had any difficulty with this aspect of the task. Full details on the warm-up procedure can be found in Flavell et al. (1993).

The task proper involved E2 engaging in three different kinds of activity: waiting, looking, and problem solving. Following Flavell et al. (1993), there were two waiting trials, which were always presented as the first and third trials in the task, and the same order of activities was employed for every child. The order of trials was: waiting, looking, waiting, and problem solving.

The two waiting trials were identical. E1 asked E2 to move to a chair, saying, “Can you just wait over there for a few minutes? I’ll tell you when we’re ready.” After a brief pause, E1 said to the child, “She’s just sitting there waiting, isn’t she?” The test question was then presented: “What about her mind right now? Is she having some thoughts or ideas, or is her mind empty of thoughts or ideas? Point to the picture that shows her mind while she’s waiting over there.” On both of the waiting trials, E2 sat quietly with her back to the child and E1. On the looking trial, E1 pointed to some pictures on the wall, saying, “I hung those pictures up the other day.” E2 responded, “Oh, I can see them.” E2 stayed with her profile turned toward the child, looking at the pictures. E1 then asked the child to select a picture of E2’s mind while she was looking at the pictures.

Between the second and third trials, E1 asked children the original familiarization questions once again to check that they had not forgotten the two mind states. On the third trial (waiting), E1 directed E2 to a chair at the opposite side of the room, away from the pictures, and instructed her just to wait once again. After this waiting trial, children were given the problem-solving trial, which involved the experimenters puzzling over a ship-in-a-bottle ornament. In the problem-solving trial, E1 said to E2, “Can you come here, because I’m ready for you to do something else now. Can you tell me how this big ship got into this little bottle?” E2 looked pensive and said, “Mmm, that’s a hard question. Give me just a minute.” The child was then asked to select a picture to represent E2’s mind at that moment.

For each of the four trials, children received a score of 0 if they chose the “empty of thoughts or ideas” bubble and 1 if they chose the “some thoughts or ideas” bubble, giving an overall score between 0 and 4.

Results

Convergent Validity of MM

Table 1 shows the descriptive statistics for all 6-month and follow-up variables. Intercorrelations between these variables are shown in Table 2. The intercorrelations between the indexes of maternal MM at 6 months (with the exception of inappropriate mind-related comments) were reported in Meins et al. (2001). A very similar pattern of correlations was found for this subset of the original sample, with 4 of 10 correlations reaching significance. Relations between mothers’ inappropriate mind-related comments and the other MM indexes were not reported in Meins et al.’s (2001) study. As Table 2 shows, none of these correlations was significant.

Temporal Continuity of MM

Table 2 shows the correlations between the 6-month indexes of MM and mothers’ mentalistic
descriptions of their children at 48 months. Mothers were more likely to describe their children in mentalistic terms if they had earlier tended to comment appropriately and refrain from commentinappropriately, on their infants’ mental states at 6 months. Mothers’ mentalistic descriptions were not, however, related to the other 6-month indexes of MM nor to maternal sensitivity. Neither early nor later indexes of MM were related to children’s BPVS–II scores. A forward regression analysis using

Table 1
Descriptive Statistics for 6-Month and Follow-Up Variables

<table>
<thead>
<tr>
<th>Variable</th>
<th>N</th>
<th>M</th>
<th>SD</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>6-month variables</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Mothers’ responses to changes in infants’ line of gaze</td>
<td>52</td>
<td>0.50</td>
<td>0.18</td>
<td>0.11–0.84</td>
</tr>
<tr>
<td>2. Mothers’ responses to infants’ object-directed activity</td>
<td>52</td>
<td>0.89</td>
<td>0.12</td>
<td>0.44–1.0</td>
</tr>
<tr>
<td>3. Mothers’ imitation of infants’ vocalizations</td>
<td>52</td>
<td>0.07</td>
<td>0.05</td>
<td>0–0.19</td>
</tr>
<tr>
<td>4. Maternal encouragement of autonomy</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Proportion</td>
<td>52</td>
<td>0.06</td>
<td>0.04</td>
<td>0–0.17</td>
</tr>
<tr>
<td>Total</td>
<td>52</td>
<td>8.39</td>
<td>5.39</td>
<td>0–22</td>
</tr>
<tr>
<td>5. Appropriate mind-related comments</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Proportion</td>
<td>52</td>
<td>0.10</td>
<td>0.06</td>
<td>0.0–0.28</td>
</tr>
<tr>
<td>Total</td>
<td>52</td>
<td>14.17</td>
<td>9.41</td>
<td>0–40</td>
</tr>
<tr>
<td>6. Inappropriate mind-related comments</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Proportion</td>
<td>52</td>
<td>0.02</td>
<td>0.02</td>
<td>0–0.09</td>
</tr>
<tr>
<td>Total</td>
<td>52</td>
<td>1.83</td>
<td>2.70</td>
<td>0–13</td>
</tr>
<tr>
<td>7. Maternal sensitivity</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. Mothers’ highest educational level</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Follow-up variables (45–55 months)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. Mothers’ mentalistic descriptions of children</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Proportion</td>
<td>51</td>
<td>0.44</td>
<td>0.23</td>
<td>0–1.0</td>
</tr>
<tr>
<td>Total</td>
<td>51</td>
<td>4.02</td>
<td>2.79</td>
<td>0–12</td>
</tr>
<tr>
<td>10. Theory of mind performance</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11. Stream of consciousness performance</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12. British Picture Vocabulary Scale II score</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2
Intercorrelations Between All Variables

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>6-month variables</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Mothers’ responses to changes in infants’ line of gaze</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.55**</td>
</tr>
<tr>
<td>2. Mothers’ responses to infants’ object-directed activity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.24</td>
</tr>
<tr>
<td>3. Mothers’ imitation of infants’ vocalizations</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.23</td>
</tr>
<tr>
<td>4. Maternal encouragement of autonomy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.48**</td>
</tr>
<tr>
<td>5. Appropriate mind-related comments</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.21</td>
</tr>
<tr>
<td>6. Inappropriate mind-related comments</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.03</td>
</tr>
<tr>
<td>7. Maternal sensitivity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.09</td>
</tr>
<tr>
<td>8. Mothers’ highest educational level</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.01</td>
</tr>
<tr>
<td>Follow-up variables (45–55 months)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. Mothers’ mentalistic descriptions of children</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.06</td>
</tr>
<tr>
<td>11. Stream of consciousness performance</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.12</td>
</tr>
<tr>
<td>12. British Picture Vocabulary Scale II score</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.03</td>
</tr>
</tbody>
</table>

Figures in italics correspond to correlations reported in Meins et al. (2001). Figures in bold correspond to correlations reported in Meins et al. (2002). In all cases correlations were derived from a subset of the original sample.

*p < .05.* *p < .01. **p < .001. All tests were two-tailed.
the six 6-month MM indexes as independent variables showed that appropriate ($t = 3.60$, $\beta = 0.42$, $p < .001$) and inappropriate ($t = -4.07$, $\beta = -0.47$, $p < .001$) mind-related comments accounted, respectively (and with opposite directions of effect), for 17% and 19% of the variance in mothers’ mentalistic descriptions at 48 months. No other 6-month MM indexes were significant predictors of later MM.

Predictive Validity of MM

Table 2 shows the intercorrelations between early and later measures of MM, ToM, and SoC performance. As the relations between early MM and ToM were the focus of Meins et al.’s (2002) study, these results are not discussed further here, except to point out that the pattern of intercorrelations in the Meins et al. (2002) study is replicated in this subset of the original sample. Two previously unreported findings are of particular interest here: (a) that, with the exception of appropriate mind-related comments, none of the 6-month MM measures was related to ToM performance; and (b) that there was no correlation between mothers’ mentalistic descriptions of their children at 48 months and the composite ToM score.

Next, we turn to early and later MM measures as predictors of SoC performance at 55 months. The relevant correlations are shown in Table 2. SoC performance was positively related to mothers’ appropriate mind-related comments at 6 months. None of the other early or later MM measures was related to SoC performance. The positive correlation between ToM and SoC performance was marginally significant ($p = .09$). Thus, there was only weak evidence for a relation between children’s understanding of basic mental states and their subsequent understanding of ongoing ideation.

Table 3 shows children’s performance on the individual trials of the SoC task. For comparison, this table also includes the numbers of children of the same age who passed the individual trials in Flavell et al.’s (1993) original study. No significant differences were found between the two studies on any of the four trials, $\chi^2(1) = 0.05$ to 2.31, $ns$. Our results thus replicate those of Flavell et al.

Divergent Validity of MM

The divergent validity of MM was addressed by comparing the predictive power of early and later measures of MM with a more general measure of mothers’ responsiveness to their children, namely, maternal sensitivity. First, we attempted to deter-

<table>
<thead>
<tr>
<th>Trial</th>
<th>Present study</th>
<th>Flavell et al. (1993; $N = 20$)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Waiting 1</td>
<td>Looking</td>
</tr>
<tr>
<td>Percentage of children</td>
<td>29%</td>
<td>35%</td>
</tr>
<tr>
<td>Attributing thoughts or</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ideas to the experimenter</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\chi^2$</td>
<td>0.05</td>
<td>2.31</td>
</tr>
</tbody>
</table>

Table 4

Results of Regression Analysis to Identify Which Early Mind-Mindedness Variables Predict Stream of Consciousness Performance

<table>
<thead>
<tr>
<th>Independent variable</th>
<th>$R^2$</th>
<th>$t$</th>
<th>$\beta$</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mothers’ responses to changes in infants’ line of gaze</td>
<td>0.02</td>
<td>-1.11</td>
<td>-0.21</td>
<td>$ns$</td>
</tr>
<tr>
<td>Mothers’ responses to infants’ object-directed activity</td>
<td>0.09</td>
<td>-2.50</td>
<td>-0.36</td>
<td>&lt;.025</td>
</tr>
<tr>
<td>Mothers’ imitation of infants’ vocalizations</td>
<td>0.01</td>
<td>-0.59</td>
<td>-0.08</td>
<td>$ns$</td>
</tr>
<tr>
<td>Maternal encouragement of autonomy</td>
<td>0.00</td>
<td>0.15</td>
<td>0.02</td>
<td>$ns$</td>
</tr>
<tr>
<td>Appropriate mind-related comments</td>
<td>0.11</td>
<td>3.30</td>
<td>0.48</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Inappropriate mind-related comments</td>
<td>0.00</td>
<td>0.16</td>
<td>0.02</td>
<td>$ns$</td>
</tr>
</tbody>
</table>
Concerning the question of the divergent validity of MM, our findings suggest that early MM (as indexed by appropriate mind-related comments) is a better predictor of SoC understanding than maternal sensitivity. These findings are thus in line with those of Meins et al. (2002), which showed appropriate mind-related comments to be a better predictor of later ToM than maternal sensitivity.

Path Analyses

To investigate the developmental pathways from early MM to children’s later understanding of mind, the data were analyzed using maximum likelihood estimation via AMOS 4.01 (Arbuckle 1999). Analyses were conducted with and without the one participant who did not complete the maternal interview at 48 months (in the former case allowing AMOS to estimate the means and intercepts). This manipulation made little difference to the results; thus, only the latter analyses are reported. Data were checked for univariate and multivariate normality and were not found to be problematic.

Developmental Pathways to ToM Performance

The first series of models explored pathways to children’s ToM performance. Three models were proposed to account for the finding that MM at 6 months related to children’s ToM performance. Maternal education and children’s BPVS–II scores were included in these analyses because they were significantly correlated with ToM performance. Fit indexes for all models are shown in Table 6.

Model 1. In this model (see Figure 1), the relation between appropriate mind-related comments at 6 months and later ToM was assumed to be indirect, functioning via mothers’ mentalistic descriptions at 48 months. There is therefore no direct pathway from appropriate mind-related comments to ToM performance. The total variance in ToM accounted for by this model was 15% ($R^2 = .15$). Fit indexes (see Table 6) indicated a poor fit. Relative effects include the following: (a) appropriate mind-related comments had a significant direct path to mothers’ mentalistic descriptions, (b) BPVS–II scores had a significant direct path to ToM, and (c) the paths from mothers’ mentalistic descriptions and maternal education to ToM were not significant.

Model 2. This model (see Figure 2) is based on the assumption that both early and later MM make independent contributions to children’s ToM perfor-
mance, with direct pathways from both MM measures to ToM. The total variance accounted for by this model was 24% ($R^2 = .24$). Fit indexes suggested a reasonable fit. Relative effects include the following: (a) appropriate mind-related comments had a significant direct path to mothers’ mentalistic descriptions, (b) appropriate mind-related comments had a significant direct path to ToM, (c) BPVS–II scores had a significant direct path to ToM, and (d) the paths from mothers’ mentalistic descriptions and maternal education to ToM were not significant. The former path (from mothers’ mentalistic descriptions to ToM) was negative and therefore reduced the total effect from appropriate mind-related comments to ToM.

Model 3. In this model (see Figure 3), early, but not later, MM is assumed to predict ToM performance. This model thus represents both later MM and children’s ToM as outcome variables of appropriate mind-related comments at 6 months. The total variance accounted for by this model was 23% ($R^2 = .23$). Fit indexes suggested a reasonable fit. Relative effects include the following: (a) appropriate mind-related comments had a significant direct path to mothers’ mentalistic descriptions, (b) appropriate mind-related comments had a significant direct path to ToM, (c) BPVS–II scores had a significant direct path to ToM, and (d) the path from maternal education to ToM was not significant.

Comparison of models. Because of the differences in degrees of freedom between Models 1 and 2 and between Models 2 and 3, it was possible to make comparisons within these two pairs on the basis of change in $\chi^2$. Model 2 was a significant improvement over Model 1, $\chi^2$diff(1) = 6.99, $p = .008$, but not over Model 3, $\chi^2$diff(1) = 0.83, $p = .36$. With respect to the fit indexes to compare Models 2 and 3, apart from standardized root mean square residual (SRMR; which suggests little difference) and Modified Expected Cross-Validation Index (MECVI), all show Model 3 to be a better fit than Model 2.

Developmental Pathways to SoC Understanding

The path analyses reported above formed the basis for a second series of analyses to model potential pathways to children’s SoC understanding.
at 55 months. The model found to be the best fit for ToM performance (Model 3) was used as the starting point for this series of analyses. Three models were proposed to investigate whether the links between early MM and later SoC understanding were direct or functioned via children’s ToM performance. Fit indexes are shown in Table 6.

**Model 4.** In this model (see Figure 4), the relation between appropriate mind-related comments at 6 months and children’s SoC understanding was assumed to be indirect, functioning via ToM performance. In this model, there is therefore no direct pathway from appropriate mind-related comments to SoC. The total variance accounted for by this model was 5% ($R^2 = .05$). Fit indexes suggested a reasonable fit. The relative effect is that the path from ToM to SoC was not significant.

**Model 5.** This model (see Figure 5) is based on the assumption that both appropriate mind-related comments and ToM performance make independent contributions to children’s SoC understanding, with direct pathways from both of these measures to SoC. The total variance accounted for by this model was 10% ($R^2 = .10$). Fit indexes suggested a reasonable fit. Relative effects include the following: (a) the path from ToM to SoC was not significant, and (b) the path from appropriate mind-related comments to SoC was not significant.

**Model 6.** In this model (see Figure 6) appropriate mind-related comments, but not ToM performance, is assumed to predict SoC understanding. This model thus represents both ToM and SoC as outcome variables of appropriate mind-related comments. The total variance accounted for by this model was 9% ($R^2 = .09$). Fit indexes suggested a reasonable fit. The relative effect is that the path from appropriate mind-related comments to SoC was significant.

**Comparison of models.** Because of the differences in degrees of freedom between Model 5 and the other two models, it was possible to make comparisons between the models. Neither of these comparisons suggested a significant improvement, but the difference between Model 5 and Model 4 was more marked, $\chi^2\text{diff}(1) = 3.12, p = .08$, than that between Model 5 and Model 6, $\chi^2\text{diff}(1) = 0.87, p = .35$. Comparison of the various fit indexes for Models 5 and 6 suggest that Model 6 is a better fit for the number of parameters to be estimated: the confidence intervals for root mean square error of approximation (RMSEA), MECVI, and SRMR are all smaller in Model 6 than in Model 5. In addition, the extra path from ToM to SoC in Model 5 only explains an extra 1.3% of the variance in SoC.

The results of these path analyses suggest that mothers’ MM at 6 months (indexed by appropriate mind-related comments) is directly related to children’s SoC understanding at 55 months, and that this relation does not function via children’s ToM performance. In neither of the models including a
direct link between ToM and SoC (Models 4 and 5) was this path significant, suggesting that there is no strong relation between children’s understanding of basic mental states and their later understanding of ongoing ideation.

Discussion

The main aim of the study reported here was to investigate the construct validity of maternal MM in the context of its value as a predictor of children’s later understanding of mind. Specifically, we expected meaningful patterns of correlation between MM measures taken at 6 months, between these variables and a measure of MM taken at 48 months, and between these early and later MM measures and children’s ToM and SoC performance. The use of path analyses allowed us to address the question of whether individual differences in MM have their effects in infancy or relatively later in development.

The first hypothesis concerned the convergent validity and temporal continuity of the construct of MM. Turning first to the interrelations between the 6-month MM measures, no strong support was found for convergent validity because only 4 of 10 correlations reached statistical significance. However, a stricter test of the convergent validity of MM is to look at how these 6-month measures relate to the later (48-month) measure of MM, namely, mothers’ mentalistic descriptions of their children. This later MM measure turned out to be positively associated with appropriate mind-related comments and negatively associated with inappropriate mind-related comments, but unrelated to any of the other infancy variables. Thus, although there was no evidence that MM at age 4 was predicted by those 6-month measures of MM that did not involve mind-related comments, strong temporal continuity was observed between early and later MM for those variables involving mothers’ explicit linguistic representations of their infants’ mental states.

The second hypothesis addressed the issue of the predictive validity of MM and entailed the prediction that both early and later MM measures would relate to a composite measure of ToM understanding. The only 6-month MM measure that correlated with children’s ToM performance was appropriate mind-related comments. The later measure of MM, mothers’ mentalistic descriptions of their children at 48 months, was not related to concurrent ToM performance.

The third hypothesis, also concerned with the predictive validity of MM, was that both early and later MM measures would predict SoC understanding at 55 months. A regression analysis using the 6-month MM measures showed SoC to be positively related to appropriate mind-related comments and negatively related to maternal responses to infants’ object-directed action. None of the other 6-month MM measures was related to SoC understanding. These variables remained the only independent predictors of SoC understanding on a further regression analysis incorporating mothers’ mentalistic descriptions at 48 months, maternal sensitivity, ToM composite scores, receptive verbal ability, and maternal education.

The fourth hypothesis concerned the divergent validity of MM. We predicted that early and later MM would be better predictors of ToM and SoC performance than maternal sensitivity. As already noted, early MM (as indexed by appropriate mind-related comments) was significantly associated with SoC understanding at 55 months, whereas maternal sensitivity showed no association with this variable. This finding is thus in line with Meins et al.’s (2002) observation that maternal sensitivity did not predict ToM once 6-month MM had been taken into account. Later MM was not related to either ToM or SoC understanding. Thus, although our hypothesis about divergent validity was supported for early MM, it was not supported for the later MM measure.

The final hypothesis concerned the relation between ToM and SoC understanding. There was only a marginally significant correlation between these variables, and a regression analysis showed that ToM was not an independent predictor of SoC understanding. We therefore found only weak support for the supposition that ToM and SoC tasks tap related aspects of children’s understanding of mind.

Finally, path analyses were used to investigate possible developmental pathways between early and later MM and children’s understanding of mind. First, three possible pathways relating early MM to ToM performance were compared. The analyses suggested that this relation was direct and did not function via later MM. We then compared three further models relating early MM to later SoC understanding. Again, the analyses suggested that the link between early MM and later SoC was direct and did not function via later ToM performance. This lack of any relation between later MM (as indexed by mothers’ mentalistic descriptions of their children) and ToM performance at 45 to 48 months is in line with the findings of Meins et al. (1998), who showed that performance on the unexpected transfer task was not significantly related to mothers’ mentalistic descriptions at age 3.
The analyses reported here therefore go some way to answering the question posed by Meins et al. (2002) regarding possible time points at which MM might have its influence on mentalizing development. The findings suggest that any influence of MM does not occur through continuity between early and later manifestations of MM. Rather, the analyses indicate that the relations between early MM and later mentalizing capacities are direct and not mediated by later MM. At first glance, the apparent lack of any mediational role of later MM is surprising, and we suggest at least two reasons why it should be interpreted with caution. First, unlike the observation-based 6-month MM measures, the later measure of MM was obtained from mothers’ descriptions of their children in an interview. The measure is therefore reflective and retrospective and may be less accurate as an indicator of MM than, for example, the on-line use of appropriate mental state language during actual mother–child interactions. It may be that the use of such interaction-based measures might result in even more impressive temporal continuity between early and later MM and in stronger correlations between later MM and children’s concurrent understanding of mind. For example, previous studies have found associations between ToM and maternal language during the negotiation of sibling disputes (Dunn et al., 1991) and book reading (Ruffman et al., 2002). One task for the future is therefore to investigate how interaction-derived measures of later MM relate to concurrent interview-based and early observation-based measures of MM, as well as to children’s concurrent mentalizing abilities.

A second possible reason for the lack of any association between later MM and the outcome measures studied here is that MM may exert its effects during the months leading up to the emergence of fully fledged representational ToM abilities. If this were the case, assessing MM concurrently with or after assessing ToM performance might fail to pick up any influence of later MM. That said, at least two findings give us reason to doubt this proposition. First, later MM was not related to SoC performance 7 months later, suggesting that any such delayed-action effect does not apply to SoC development. Second, Meins et al. (1998) found no relation between maternal MM at 36 months and children’s performance on the unexpected transfer task at 48 months. It therefore seems unlikely that the apparent lack of influence of later MM can be explained simply in terms of our having measured it at an inappropriate age.

How can one account for the finding that it is early rather than later MM that is most important as a predictor of later mentalizing abilities? Meins et al. (2002) have suggested that mothers’ appropriate comments on their infants’ mental states might provide a linguistic and conceptual scaffold within which infants can begin to understand how mental states determine behavior. Need we, on the basis of our present findings, claim that this process begins as early as 6 months? Many readers might find it implausible that such young infants could be decoding mothers’ mentalistic comments and using them as a source of information on their own mental states. Rather, one might wish to claim that it is mothers’ use of appropriate mind-related comments during the first year or so of life, rather than as early as 6 months, that has the crucial influence on children’s subsequent understanding of mind. Specifically, one would expect that mothers who comment appropriately on their infants’ mental processes at 6 months will continue to do so at 12 and 18 months when their infants are actively acquiring language, including mental state language (Bretherton, McNew, & Beeghly-Smith, 1981). Thus, it may be the continuity of MM discourse during infancy, rather than in the preschool years, that accounts for the observed relations with later mentalizing abilities. However, our results show that it is only a mother’s use of mental state language that matches her infant’s current state of mind that relates to later mentalizing, because inappropriate mind-related comments showed no association with ToM or SoC understanding. We can therefore suggest that, at the very least, appropriate mind-related comments at 6 months index the beginnings of mothers’ and infants’ joint attention to mind, in which mothers’ appropriate labeling of their infants’ current mental states helps draw infants’ attention to the existence (and perhaps functional significance) of mental states and processes. In due course, and perhaps at a slightly later stage in development, this kind of interaction presents infants with opportunities to integrate subjective information on their mental states with an external linguistic comment on the behavior that results from these states.

This view is in line with that of Harris (1996; Harris & Leevers, 2000), who has suggested that observed links between conversational language and ToM understanding can be explained in terms of preschool children integrating external and internal information about the relations among beliefs, desires, and behavior. The present findings suggest, however, that this process may begin long before
infants acquire the linguistic skills to be able to engage in conversation for themselves. It may be that, rather than building on rudimentary ToM abilities that are already present (as suggested by Harris), early exposure to MM language has its effects even earlier, before children have acquired any understanding of mind. That is, it may be one of the routes by which infants become aware of their own and others’ mental states in the first place.

Related to this argument is Flavell et al.’s (1993) suggestion that the lag in acquiring an understanding of the SoC (relative to ToM) is at least partly a result of the fact that people tend not to comment on unobservable mental processes. In fact, research on MM has shown that certain mothers do comment on such processes from the earliest months of life. For example, Meins et al. (2001) reported that mothers comment that their 6-month-old infants are thinking, making decisions, remembering, and even teasing and joking. We suggest that these comments play a causal role in helping children to acquire an understanding of ongoing ideation: the interplay of thoughts, beliefs and ideas that makes up the SoC. If our proposal is correct, the fact that mothers engage in such discourse from the earliest months of life (Meins et al., 2001) means that, by the time they reach their fifth year, children will already have had plenty of opportunity to reflect on how the SoC continues in the absence of any external activity.

Although the positive relation between appropriate mind-related comments and children’s subsequent SoC understanding was predicted, one unexpected finding was the negative relation between one of the other indexes of early MM (maternal responsiveness to infant’s object-directed action) and performance on the SoC task. How should we square this finding with the evidence from the 6-month assessment, where responsiveness to object-directed action was positively associated with mothers’ use of appropriate mind-related comments? One possibility is that this negative relation is a function of the age at which the measures were taken. At 6 months, when the infant’s main cognitive challenge is arguably to master the manipulation of objects, a mother’s focus on such activity might indeed be taken as indicative of MM. However, in later years, a continuing focus on object-directed activity (rather than on language, emotions, or more general exploration of the social and physical worlds) would not fit with our characterization of the MM mother. A challenge for future research, then, is to gather evidence on the temporal continuity of mothers’ responsiveness to object-directed action, with a view to determining whether such responsiveness at different ages has differing developmental outcomes.

The present study also allowed us to perform some important tests of Flavell et al.’s (1993) original findings on children’s SoC understanding and to situate these findings within the context of children’s general ToM development. First, we found similar patterns of failure and success in our sample compared with Flavell et al.’s. Second, our inclusion of ToM data meant that we were able to investigate the relation between SoC and ToM. The lack of any relation between these variables is at first glance surprising and suggests that the SoC task taps different capacities from those assessed by the ToM battery. Such a view is in line with recent findings that ToM abilities might not be as homogeneous as previously thought. For example, a number of recent studies have failed to find high levels of internal consistency for composite ToM measures (Astington & Jenkins, 1999; Hughes & Dunn, 1998). Indeed, the view that the development of ToM and SoC abilities is strongly influenced by varieties of interactional experience gives one strong reasons for predicting precisely such heterogeneity.

What, then, can be said about the relation between ToM and SoC understanding? Our suggestion is that a rudimentary grasp of belief–desire psychology (as assessed by standard ToM tasks) is a necessary, but not sufficient, condition for acknowledging the SoC. In determining what else might be important in this regard, it is worth bearing in mind that the SoC, at least in normal adults, is a primarily verbal phenomenon (Carruthers, in press; Hurlburt, 1990). One might therefore expect that a child’s developing understanding of the SoC will be influenced by the same factors that affect the emergence of verbal thought. There are at least two reasons for believing that this might be the case. First, if thoughts and ideas (the terms used in the SoC task) are conceived as essentially verbal phenomena, children are unlikely to attribute them to a silent other before they have acquired inner speech for themselves. Second, the acquisition of inner speech may in itself provide children with a tool for reflecting on their own mental processes. According to Vygotsky (1934/1987), the acquisition of verbal thought or inner speech is the result of a gradual process of internalization of social speech through the intermediary stage of private speech. This raises the possibility that children’s acknowledgment of ongoing ideation may relate to their use of task-relevant private speech, and we are currently testing this hypothesis. Another possibility is that SoC understanding may relate to the emergence of
autobiographical memory, a further context within which children might reflect on their own ongoing ideation.

Two further points need to be made in conclusion. It is important to note that what we have portrayed as a social-environmental influence on mentalizing development may in fact be no such thing, and that the associations between early maternal MM and children’s mentalizing abilities may merely reflect genetically determined differences in ToM competence transmitted from mother to child. Because the mothers in our study were all the biological parents of the participating children, our data cannot speak to this question. Instead, while awaiting the outcome of future research on the genetic components of ToM, we can note that comparisons of MM between biological and adoptive mothers might be a useful avenue for exploring this question. Another important priority for future research is to determine whether MM is specific to a particular mother–child relationship or whether it represents a general trait in the mother. Evidence for the former would strengthen our case that MM has a true social-environmental influence on mentalizing development rather than merely reflecting genetically determined individual differences in ToM.

Finally, we turn again to the issue of convergent validity with which we began our investigation. Although strong evidence was found for temporal continuity between MM measures that involved mothers’ explicit linguistic representations of their infants’ mental states, it is still not possible to draw any firm conclusions regarding the relation between these measures of MM and those that do not tap mothers’ explicit mentalistic representations of their infants. This is at least in part because we have not yet been able to supplement our knowledge about these latter nonexplicit MM measures with equivalent measures in the preschool years. For example, it would be useful to know how a nonexplicit MM measure such as encouragement of autonomy might, when measured at a later stage in development, relate to children’s emerging mentalizing abilities. Such a research program would allow us to answer more fully the question of whether all the maternal behaviors that we have designated as MM, and not just those that involve explicit mentalistic representations, are important for mentalizing development. That said, the lack of any clear influence of nonexplicit MM measures would suggest that it is mothers’ explicit and appropriate talk about their infants’ mental states that is of most importance in facilitating children’s developing understanding of mind.

References


Harris, P. L. (1996). Desires, beliefs, and language. In P. Carruthers, & P. K. Smith (Eds.), *Theories of theories of


