Developmental Psychology
2008, Vol. 44, No. 4, 1055–1068

0012-1649/08/$12.00 DOI: 10.1037/0012-1649.44.4.1055

Theory-of-Mind Development Influences Suggestibility and Source Monitoring

Alexandra Bright-Paul and Christopher Jarrold
University of Bristol

Daniel B. Wright
University of Sussex

According to the mental-state reasoning model of suggestibility, 2 components of theory of mind mediate reductions in suggestibility across the preschool years. The authors examined whether theory-of-mind performance may be legitimately separated into 2 components and explored the memory processes underlying the associations between theory of mind and suggestibility, independent of verbal ability. Children 3 to 6 years old completed 6 theory-of-mind tasks and a postevent misinformation procedure. Contrary to the model's prediction, a single latent theory-of-mind factor emerged, suggesting a single-component rather than a dual-component conceptualization of theory-of-mind performance. This factor provided statistical justification for computing a single composite theory-of-mind score. Improvements in theory of mind predicted reductions in suggestibility, independent of verbal ability (Study 1, $n = 72$). Furthermore, once attribution biases were controlled (Study 2, $n = 45$), there was also a positive relationship between theory of mind and source memory, but not recognition performance. The findings suggest a substantial, and possibly causal, association between theory-of-mind development and resistance to suggestion, driven specifically by improvements in source monitoring.

Keywords: children's eyewitness suggestibility, theory of mind, source monitoring

Supplemental materials: http://dx.doi.org/10.1037/0012-1649.44.4.1055.supp

One of the most frequently documented phenomena in the eyewitness memory literature is the postevent misinformation effect. Misleading information presented after a witnessed event may be unwittingly incorporated into subsequent recalled reports of the event, a phenomenon commonly referred to as *suggestibility* (Loftus, 1979). The effect has been demonstrated under a variety of test conditions, including free recall, directed recall, recognition, and source-directed tasks, and is evident in adults and children alike (for reviews, see Ayers & Reder, 1998; Bruck & Ceci, 1999). Nevertheless, the majority of research indicates that children are more suggestible than adults, and the younger the child, the more susceptible they tend to be. In particular, there is a marked increase in resistance to suggestion between the ages of 3 and 6 years (Ceci & Bruck, 1993). Recently, the literature has shifted from merely describing these age differences to trying to unravel the multiple social and cognitive mechanisms that underpin them (Bruck & Melnyk, 2004; Quas, Qin, Schaaf, & Goodman, 1997). One aspect of cognitive development that has received particular attention for explaining these variations is the growth in understanding of the representational nature of mind, or *theory of mind*.

Alexandra Bright-Paul and Christopher Jarrold, Department of Experimental Psychology, University of Bristol, Bristol, England; Daniel B. Wright, Department of Psychology, University of Sussex, Brighton, England.

We thank Blaise, Sefton Park, and Shirehampton Infants and Nursery Schools in Bristol for their help with the research. The research was conducted as work toward a doctoral thesis at the University of Bristol and was funded by an Economic and Social Research Council Postgraduate Research Studentship. Some of these findings were presented at the Society for Research in Child Development's 2001 biennial meeting, Minneapolis, Minnesota.

Correspondence concerning this article should be addressed to Alexandra Bright-Paul, Department of Experimental Psychology, University of Bristol, 12a Priory Road, Bristol, BS8 1TU England. E-mail: a.m.bright-paul@bristol.ac.uk

Theory of Mind

An understanding of mind implies that an individual is able to "impute mental states to themselves and to others" (Astington, 1993, p. 4). It is seen as a theory because mental states are not directly observable and can be used to make predictions about the behavior of others. The benchmark test for demonstrating this understanding is the *false-belief task* (Dennett, 1978). Although modifications to the procedure may result in early success (P. Mitchell & Lacohee, 1991), children are typically unable to attribute false beliefs to themselves or to others until 4 years of age (see Wellman, Cross, & Watson, 2001). At approximately the same age, children understand that the appearance of an object is a representation, dependent on perspective (the *appearance–reality distinction*), although this understanding may appear earlier in a pretend-play context, because 3-year-olds are able to identify the real and atypical functions of objects during pretence (*pretend–real distinction*; Flavell, Flavell, & Green, 1987). According to Perner (1991), success across these tasks at age 4 years indicates a major conceptual shift in representational understanding.

An alternative view is that success on these and other tasks reflects a growth in understanding that knowledge is dependent on informational access (Wimmer, Hogrefe, & Sodian, 1988). Although 3-year-olds associate seeing with knowledge, it is not until

around age 4 that children are able to explicitly reflect on *how* they, or another person, acquired knowledge and make explicit connections between knowledge states and the informational conditions (e.g., perceptions) from which they originate. This ability is measured by *origins-of-belief tasks* in which children identify the means by which they acquire knowledge (e.g., "I know because I saw it/heard it"). A further refinement appears between ages 4 and 5, when children appreciate that different modes of access to information (e.g., seeing, touching) inform different aspects of knowledge (e.g., color, texture), a skill referred to as *aspectuality* (O'Neill, Astington, & Flavell, 1992). Children who fail aspectuality tasks are unable to predict whether they need to look or to feel pairs of objects that differ either visually or by texture or weight in order to identify them. Finally, at around 5 to 6 years of age, children begin to appreciate that another's knowledge may also be drawn from mental inferences (Sodian & Wimmer, 1987) and not solely from direct perceptual access. Before this age, children exhibit *inference neglect* and do not acknowledge that another person is able to know the contents of a box by merely inferring this information, rather than by seeing it directly (Varouxaki, Freeman, Peters, & Lewis, 1999).

Theory of Mind and Suggestibility Associations

Given that the emergence of theory of mind and improvements in resistance to suggestion coincide between ages 3 and 5 years, it seems plausible that there may be a link between the two developments. A number of empirical studies have addressed this possibility by examining the relationship between theory of mind and suggestibility. For example, Welch-Ross, Diecidue, and Miller (1997) showed 3- to 5-year-olds a visual scene (event) and misled them about key details (misinformation) before posing yes–no recognition questions about the event (test). Children also completed pretend–real, appearance–reality, and false-belief tasks, from which a composite theory-of-mind score was computed. The amount of misleading detail falsely recognized as having occurred in the original scene was negatively associated with theory-of-mind performance, accounting for 25% of the variance, after controlling for chronological age and memory for neutral items. Similarly, Templeton and Wilcox (2000) reported that other false-belief success accounted for 25% of the variance in the misinformation effect in their study. The extent of the association may be contingent on whether the memory for the original event is strong (Welch-Ross, 1999), whether theory of mind is measured by own or other false-belief tasks (Quas & Schaff, 2002; Thomsen & Bernstein, 2005), and whether the child is questioned under social pressure (Scullin & Bonner, 2006).

Despite the growing number of studies reporting an association between theory of mind and suggestibility, there has been no clear consensus regarding explanations of the mechanisms that underpin the relationship. To date, explanations for the link have tended to be speculative and nonspecific. For example, Templeton and Wilcox (2000) suggested similarities in the skills required for resisting suggestion and for understanding mind. Both require evaluation of two representations held simultaneously (event vs. misinformation; reality vs. false belief) and formation of a belief about what actually happened. Others have proposed that false-belief understanding allows children to adopt more sophisticated strategies for handling interviewer questions (Scullin & Bonner, 2006). In particular, if children understand that the questioner can entertain a false belief about a given situation, then this provides them with the necessary skills to resist suggestion (see Bruck & Melnyk, 2004).

In contrast, Welch-Ross (2000) offered a precise framework for interpreting the memory processes involved in connection between suggestibility and specific aspects of theory-of-mind development. According to her mental-state reasoning model of suggestibility (Welch-Ross, 2000), two distinct components of theory of mind contribute to reductions in suggestibility in separate and specific ways. First, the ability to reason about conflicting mental representations, measured by success on pretend–real, false-belief, and apparent–real tasks, is said to aid resistance to suggestion by decreasing the likelihood of overwriting of the event trace with the misinformation (see Loftus, Miller, & Burns, 1978). Welch-Ross suggested that when children who fail these tasks are confronted with two conflicting representations of an event (i.e., event and misinformation), they are likely to overwrite the original memory with the misinformation, because they have difficulty simultaneously considering multiple contradictory representations (see also Thomsen & Bernstein, 2005).

The second component of theory-of-mind development that Welch-Ross (2000) suggested has an influence on suggestibility is the emergence of understanding of the connection between origins of experience and knowledge. This growth in understanding is said to underpin the meta-cognitive skills necessary for accurate source monitoring, which in turn leads to reductions in suggestibility. Source monitoring refers to the ability to identify the source or context from which memories are derived (see the source-monitoring framework; Johnson, Hashtroudi, & Lindsay, 1993). In the context of the suggestibility paradigm, even when the content of the event and misinformation phases are well remembered, the origins or source of these memories may be misidentified, resulting in suggestibility despite the coexistence of both event and misinformation traces (Lindsay & Johnson, 1989). Although the mental-state reasoning model provides a promising framework for understanding the mechanisms underpinning associations between theory of mind and suggestibility, a number of issues would benefit from clarification.

Examination of Theory-of-Mind Components

One issue is whether Welch-Ross's (2000) division of theory of mind into two distinct components (conflicting mental representations and origins of knowledge) is viable and whether each component makes separate and specific contributions to variations in suggestibility. Certainly, the theory-of-mind literature has not made such a clear distinction between mutually exclusive features of representational understanding. Indeed, the suggestion has been that the skills tapped by these types of tasks are inherently linked. For example, Perner (1991) suggested that representational understanding, measured by false-belief success, is required for understanding the importance of informational access in knowledge formation. Alternatively, Wimmer and Gschaider (2000) claimed that mastery of the false-belief task is grounded on the understanding of the causal relationship between perceptual or inferential information and resulting epistemological states. Furthermore, the suggested specific links between these so-called separate aspects of theory of mind and particular types of memory problem are not

necessarily supported empirically. For example, Naito (2003) reported associations between source monitoring and false-belief performance, even though false belief is said to measure the ability to reason about conflicting mental representations. Furthermore, in some of her analysis, Welch-Ross (1999) combined both aspects of theory of mind within a single composite score, despite the assertion that these components are measuring separate aspects of theory of mind that make distinct contributions to suggestibility.

One way forward is to use a technique similar to factor analysis to examine whether a two-component approach is supported statistically. Within the present study, this type of approach would identify the number of latent factors underlying performance on individual theory-of-mind tasks and the manner in which these individual tasks should be grouped together for a composite measure or measures. Two factors may support Welch-Ross's (2000) dual-component conceptualization. However, if performance on the individual tasks loads onto one factor, this would suggest that there is some underlying singular feature of development that is tapped by, and hence common to, all individual tasks, thus providing justification for summing performance across all tasks into a single composite score. Indeed, a single factor would fit well with a meta-analysis of several theory-of-mind tasks that demonstrated a consistent progression in mental-state understanding across a variety of tasks (Wellman & Liu, 2004).

Because theory-of-mind tasks are typically scored dichotomously (pass or fail), in the present research, we used latent trait analysis (Bartholomew, Steele, Moustaki, & Galbraith, 2002). This technique, analogous to factor analysis, allows continuous latent factors to be drawn from several binary manifest variables and allows individual task response patterns to be weighted accordingly.

Memory Processes Involved in the Association

A second issue is why theory of mind and suggestibility are related. In particular, what are the memory processes that underpin the association? In the present research, we were specifically interested in whether source monitoring is truly implicated in the relationship, as suggested by Welch-Ross (2000).

Certainly, several lines of research suggest that it is plausible that source monitoring may mediate the association between suggestibility and theory of mind. Reductions in suggestibility are accompanied by, and indeed are highly correlated with, improvements in source monitoring between the ages of 3 and 5 years (Giles, Gopnik, & Heyman, 2002). Moreover, alerting children of this age to evaluate the sources of information effectively helps them to resist suggestion (Bright-Paul, Jarrold, & Wright, 2005; Thierry & Spence, 2002). The age at which identification of external sources improves and contributes to resistance to suggestion (Poole & Lindsay, 1995) corresponds with the age at which theory of mind emerges. Indeed, theory-of-mind success is strongly associated with source-monitoring accuracy (Naito, 2003) and with episodic memory formation (Perner & Ruffman, 1995). Although findings are mixed (Farrant, Blades, & Boucher, 1998), some evidence has shown that individuals with autism, who typically show theory-of-mind deficits, are poorer than typically developing children at certain types of memory tasks that require source-monitoring skills (Russell & Jarrold, 1999).

Certainly, this combined evidence is suggestive of a link between theory of mind and suggestibility, mediated by source monitoring. However, Welch-Ross's (1999, 2000; Welch-Ross et al., 1997) findings, which she used to support her model, do not exclude the possibility that the link between theory of mind and suggestibility is driven by old–new recognition failures rather than by source monitoring per se. Old–new recognition refers to the ability to simply identify whether a target occurred, in contrast to source monitoring, which refers to the ability to also identify the origins of the target, having established that it occurred. There are several reasons why Welch-Ross's findings do not provide definitive evidence for a specific link among source monitoring, theory of mind, and suggestibility. First, Welch-Ross (1999) reported that when recognition for nonmisinformed items was controlled for, the association between theory of mind and suggestibility was weak and nonsignificant. Second, the use of yes–no recognition questions to probe memory (e.g., "Did Sally eat eggs for breakfast?") does not allow us to differentiate between recognition and source-monitoring failures, because children could have answered the question incorrectly because they had forgotten the detail rather than the source of the information. Third, although Welch-Ross (2000) did report one study in which source monitoring was measured by a task that explicitly probed for source (e.g., "Did you *see* Sally fall from the jungle-gym or did Kristin *tell you* Sally fell from the jungle-gym?"), the relationship between memory errors and origins-of-knowledge performance also extended to recognition performance measured by a separate set of forced-choice questions (e.g., "Did John wear Big Bird pajamas or did John wear Mickey Mouse pajamas?").

Taken together, this evidence does not definitively support the assertion that suggestibility is associated specifically with source monitoring, and the link may simply be driven by improvements in old–new recognition. Errors on Welch-Ross's (2000) forced-choice recognition task could conceivably have been driven either by a failure to remember that the item had occurred, irrespective of source, resulting in misses or false alarms (i.e., recognition), or by the child remembering the information but attributing it to the wrong source (i.e., source monitoring). One method for separating these two types of error is to measure suggestibility (misattributions from misinformation to event phase), source monitoring (correct identification of the source of old items), and old–new recognition (identification of items as old, regardless of source) from performance on a single source-directed task (see, e.g., Ackil & Zaragoza, 1995). By examining the associations between these separate measures and theory-of-mind performance, the present research aimed to establish whether a relationship between suggestibility and theory of mind is underpinned principally by source monitoring or more generally by old–new recognition skills.

Controlling for Verbal Ability

A third issue is controlling for verbal ability when analyzing the association between theory of mind and suggestibility. In all previous research, chronological age was partialed out in the analysis to reveal the unique association between theory of mind and suggestibility. However, we suggest that the more appropriate control is verbal ability, because it is not only a more valid index of development, but it is also highly associated with theory-of-mind performance (Sparrevohn & Howie, 1995; Yirmiya, Erel,

Shaked, & Solomonica-Levi, 1998). Arguably, this is because many theory-of-mind tasks have a linguistic component. However, this does not imply that verbal ability is simply another measure of theory of mind or, indeed, that theory of mind is a prerequisite for success on verbal ability tasks. Indeed, verbal ability, measured here by the British Picture Vocabulary Scale (Dunn, Dunn, & Whetton, 1997), is a measure of receptive language that does not require representational skills and is therefore distinct from theory of mind in this respect. Nevertheless, it is plausible that apparent associations between theory of mind and suggestibility may be driven by verbal ability, and this needs to be clarified.

Summary of the Present Research

To summarize, the present research had three aims. The first aim was to statistically examine the dual-component view of theory-of-mind performance outlined by the mental-state reasoning model of suggestibility. The second aim was to investigate the possible role of memory for source and old–new recognition in mediating the relationship between theory of mind and suggestibility. Finally, the third aim was to examine whether associations between theory of mind and suggestibility were maintained after controlling for verbal mental age.

In two studies, 3- to 6-year-olds watched an event and experienced misinformation before allocating target items to one of four possible sources at test (event, misinformation, both event and misinformation, new). From performance on this task, we calculated measures of suggestibility (misattributions from misinformation to event phase), source monitoring (correct identification of the source of old items), and old–new recognition (identification of items as old, regardless of source). Children also completed a battery of six standard theory-of-mind tasks, comprising measures of conflicting mental representation reasoning (pretend–real, own false belief, other's false belief) and origins of knowledge (origins of belief, aspectuality, and inference neglect), according to Welch-Ross's definition.[1] A latent trait technique determined whether a theory-of-mind score or scores should be based on multiple or single components. An analysis of the associations between the theory-of-mind composite score or scores and the three measures drawn from the eyewitness task assessed the extent to which theory mediated specific variations in suggestibility, source monitoring, and old–new recognition, independent of chronological and verbal mental age. We also examined whether theory-of-mind understanding was associated with recognition under nonmisleading conditions.

Study 1

Method

Participants

Participants were 72 children (33 male, 39 female) between 3 and 6 years old (M = 59.53 months, SD = 10.92, range = 42–77 months) who were recruited from a state primary school in Bristol, England. The school was selected to ensure that participants were recruited from areas of average socioeconomic status, although details of parents' education, occupation, and income were not recorded. The native language of all participants was English. There was a uniform distribution of numbers of children between 3.5 and 4.5 years old (n = 24), between 4.5 and 5.5 years old (n = 24), and between 5.5 and 6.5 years old (n = 24). The verbal mental age of participants, measured by the British Picture Vocabulary Scale (Dunn et al., 1997), ranged between 25 and 113 months (M = 68.35, SD = 20.76).

Design

Participants were divided equally into misinformed (mean verbal mental age = 67.03, SD = 21.82, n = 36) and not misinformed groups (mean verbal mental age = 69.61, SD = 19.87, n = 36). Within each group, participants were further subdivided into three film groups for counterbalancing purposes (see later description). Children were allocated to groups on the basis of chronological age to ensure a similar age range across groups (Group 1: mean verbal mental age = 67.08, SD = 21.97; Group 2: mean verbal mental age = 68.42, SD = 21.10; Group 3: mean verbal mental age = 69.54, SD = 19.97).

The three dependent measures for the misinformed group were measures of recognition, suggestibility, and source monitoring (see the Results section for calculations). The dependent measure for the nonmisinformed group was a recognition measure. Associations between memory and a composite theory-of-mind score or scores were examined, after partialing out chronological and verbal mental age.

Eyewitness Memory Procedure

Misinformed group. Children in the misinformed group undertook a three-stage postevent misinformation procedure, with materials previously tested for suitability for this age range (Bright-Paul et al., 2005). In the event phase, groups of 6 to 12 participants watched an 8-min event "film" about a theft, which was depicted on 52 slides with accompanying narrative. Twelve target items, embedded within an action, were included within the film (e.g., "Sara ate a sandwich"), each of which was represented on 2 consecutive slides. All targets were nouns typically acquired before age 36 months, according to objective age-of-acquisition norms (Morrison, Chappell, & Ellis, 1997) to minimize variations in linguistic competency influencing performance. Targets were counterbalanced across the phases of the procedure, resulting in three versions of the event, misinformation, and items presented at test (e.g., "Sara ate a sandwich/a cake/a banana").

In the misinformation phase, 1 day after the event, the experimenter read a misleading story to the same children. The story was structurally identical to the film, except that 6 of the 12 target items were changed. For example, one of the characters ate a sandwich in the film, but was described as eating a cake in the corresponding narrative. The remaining 6 target items were experienced in both the film and the story.

Immediately after the misinformation phase, participants completed a short filler task and then began the test phase. Children

[1] The expected age of success for these tasks ranged from 3 to 6 years. This range extended beyond that used by Welch-Ross (1999, 2000) but included aspects of theory-of-mind development (i.e., aspectuality and inference neglect) that Welch-Ross considered might be relevant to her model. The aim of this extended range of tasks was to provide enough variation to optimize potential associations with memory measures.

were asked to identify the source of presented targets with a source-directed forced-choice task in which they posted target items, depicted on cards, into one of four posting boxes according to source. As a visual age-appropriate analogy to adult procedures (e.g., Lindsay & Johnson, 1989), this task has been used effectively with children as young as 3 and 4 years old and is less likely to elicit a *yes* bias than other verbal versions typically used with older children (Bright-Paul, Jarrold, & Wright, 2005). The procedure also has the advantage of minimizing the influence of the quality of narrative skills and heuristic judgments at retrieval.

The four possible sources, conveyed simultaneously by pictures on each of the posting boxes, were *both event and misinformation, event only, misinformation only,* and *new*. These sources were referred to as "both film and story," "film only," "story only," and "neither" to participants. The pictures on the four posting boxes were (a) a television representing the film only option, (b) a book representing the story only option, (c) a television and a book representing the both film and story option, and (d) a large red cross representing the neither option. The target cues were pictures, drawn from a colored version of Snodgrass and Vanderwart's (1980) picture set. The use of pictures was aimed to further reduce the likelihood of linguistic competency influencing performance. At test, 24 targets were presented (4 practice and 20 test). These were 6 target items experienced in both event and misinformation phases, 6 targets experienced in the event only, 6 in the misinformation only, and 6 new items that were not experienced in either medium.

A training phase preceded the test proper, in which the purpose of the boxes was explained and participants pointed to boxes representing each source category when prompted by the experimenter (see Bright-Paul et al., 2005). In a subsequent practice phase, with 4 targets, participants selected a single target card and confirmed that they knew the object that the picture depicted (e.g., cake). The experimenter then read out the relevant corresponding phrase (e.g., "Sara ate a cake"; see Appendix A in the supplemental material for list of test phrases), and the child posted the target into one of the boxes according to source. The experimenter provided feedback, and the child was asked to post the target in the correct box, if the incorrect box was initially chosen. In the test proper, children posted 20 targets (5 of each possible source) into the source boxes, without feedback.

Not misinformed group. The materials and procedure for the not misinformed group were identical to those for the misinformed group, except for the following modifications. First, at the misinformation story stage, children were not misinformed but listened to a neutral, unrelated story ("Mr. Wrong"). Second, at test only two posting boxes were presented (film only and new), because none of the targets were witnessed at the story stage. The presented test targets comprised 10 targets experienced in the film and 10 novel targets. These targets were the same 24 target items (4 practice, 20 test) as those used for the misinformed group.

Theory-of-Mind Composite Measure

In a separate session, children completed six theory-of-mind tasks within 3 weeks of the eyewitness memory task. The expected age of success of tasks ranged between 3 and 6 years. The tasks are presented next in order of difficulty according to previous research, starting with the least difficult. Each task was scored dichotomously (pass–fail), and one point was scored for each task passed.[2]

Pretend–real. The pretend–real procedure (Welch-Ross et al., 1997) assessed ability to form two mental representations of a single object under conditions of pretence. Eight objects were presented, four of which were real (plate, cup, hairbrush, shell) and four of which represented an object (i.e., pretend objects: plastic lemon, wooden apple, plastic jam tart, ceramic egg). After the child viewed and handled an object, the experimenter stated, "Right now this looks like (e.g., an apple)," and then asked, "Is it a pretend apple or a real apple?" Participants passed the task if all eight trials were answered correctly.

False belief (other). Third-person false belief was measured with the Sally-Anne task (Baron-Cohen, Leslie, & Frith, 1985). The false-belief question was "Where will Sally first look for her marble?" and was followed by a control reality question, "Where is the marble really?" It was also followed by a control memory question, "Where did she put it in the beginning?" None of the children failed the control questions; therefore, data from all participants were included in the analysis.

False belief (own). The deceptive box procedure measured difficulties in acknowledging one's own prior false belief (Astington & Gopnik, 1988). The experimenter presented a well-known confectionery tube (Smarties candy), and the child was asked to predict the contents ("What do you think is inside the tube?"). After revealing the true contents (two crayons), the experimenter asked, "What did you think was inside the tube *before* it was opened?" This question was followed by a control reality question, "What is really inside?" Once again, none of the children failed the control questions; therefore, data from all participants were included in the analysis.

Origins of belief. The origins-of-belief task (O'Neill & Gopnik, 1991) measured children's understanding of the link between access to perceptual information and knowledge. For each of the six test trials (two each of see, tell, and no-information trials), an object (clock, monkey, sheep, elephant, lollipop, house) was placed inside a box. Children (a) watched the experimenter place the object in the box (see trials), (b) covered their eyes while the experimenter told them what she was placing inside (tell trials), or (c) covered their eyes without any information from the experimenter (no-information trials). Children were asked the following questions: "Do you know what's inside the box?" "What is it?" They were then asked the test question: "*How* do you know that?" (or "Why don't you know that?" for no-information trials). If the child was reticent to respond, a forced-choice prompt was provided: "Did you see me put it inside, or did I tell you?" (see O'Neill & Gopnik, 1991). Children passed the origins-of-belief task if all six trials were answered correctly, regardless of whether they responded to the open or forced-choice question.

Aspectuality. The aspectuality task (O'Neill et al., 1992) measured understanding of the link between perceptual knowledge and current beliefs, on the basis of inputs of a specific modality. For

[2] Note that there was a discrepancy in the number of trials for each of the individual tasks. Because we scored each task in a binary fashion (pass–fail), this provided the necessary consistency across tasks. If researchers were to use metric measures in the future, it would be preferable to have the same number of trials for each of the tasks.

Table 1
Eyewitness Memory: Descriptive Statistics

	Study 1			Study 2					
				Short event test			Long event test		
Measure	*M*	*SD*	Range	*M*	*SD*	Range	*M*	*SD*	Range
Recognition	2.55	2.06	0–8	2.48	1.78	0–7	2.95	2.09	0–8
Suggestibility	2.08	1.66	0–5	3.08	1.41	0–5	2.40	1.35	0–4
Source monitoring	0.67	0.21	0.33–1	0.47	0.19	0.14–0.93	0.52	0.23	0.20–1

each trial, a pair of objects was presented that differed along only one perceptual dimension. The object pairs were either (a) visually identical but dissimilar by touch for the two see trials (soft and hard sponges, heavy and light boxes) or (b) identical by touch and visually dissimilar for the two touch trials (red and blue toothbrushes, pink and yellow cards). Before the trials, children examined how the pairs of objects differed. The experimenter then showed the children a single object pair and asked them to cover their eyes while placing one of the items into an open-ended box covered with felt. The test question was "To find out for sure which (e.g., toothbrush) is in the box, what would you have to do? Would you have to look or feel inside the box?" After responding, children verified their response by placing their hand inside the box or lifting up the green felt to see inside. Children passed the task if they responded to all four trails correctly.

Inference neglect. The inference neglect task assessed the understanding that others may gain access to knowledge through mental inference. Even when children are initially able to use inferential information themselves, they may not extend this ability to others, a concept Sodian and Wimmer (1987) dubbed *inference neglect*. For each trial, paired objects that differed only by color (e.g., red and blue pens, yellow and blue beads, orange and yellow balls, blue and orange crayons) were placed in two boxes (one object in each box). One box was said to belong to the participant and the other to a doll (from Varouxaki et al., 1999). For the first trial, both the children and the doll "observed" an object being placed in their own box, but on the following trials, (a) only the children, (b) only the doll, or (c) neither the doll nor the children were able to "see" the placement of the item in their own box. Neither the doll nor the children were permitted to observe the object that was placed in the other's box on any of the trials. The children were first probed about their own ability to make links between inferences and knowledge: "Do you know what color (e.g., pen) is inside dolly's box?" and "What color is it?" After being reminded of whether the doll had observed placement in her own box, the children were asked the test questions: "Does dolly know what color pen is in your box?" and "What color does dolly think is in your box?" These questions assessed the children's ability to extend knowledge of inference to another person. The two critical trials were the trials in which both the doll and the children saw the object and the trials in which only the doll saw the object, because the children were required to identify whether another person could infer the color of an object from the color of the other item in the pair. The doll-only trial ensured that the children were not merely extending their own knowledge to the other person. Each trial was passed if the children acknowledged that the doll could predict the color of the object in the children's box by merely viewing the contents of her own box (and if they also correctly named the color that the doll would infer on trials in which both the doll and the children saw the object). An overall pass was obtained by passing both critical trials.

Results

Eyewitness Memory Performance

Data were collapsed across film groups because total correct attributions, measured by the sum of the number of targets allocated to their correct source (possible range = 0–20), did not vary significantly between film groups in either condition: misinformed group (Film 1: $M = 12.75$, $SD = 4.02$, $n = 12$; Film 2: $M = 13.67$, $SD = 4.03$; Film 3: $M = 12.75$, $SD = 3.33$), $F(2, 33) = 0.231$, $p = .80$; not misinformed group (Film 1: $M = 17.42$, $SD = 2.07$; Film 2: $M = 16.83$, $SD = 2.29$; Film 3: $M = 18.00$, $SD = 1.65$), $F(2, 33) = 1.00$, $p = .38$.

Three measures of performance were calculated from the eyewitness test for the misinformed group. Recognition was the sum of misses and false alarms (range = 0–20), creating a score in which a low value corresponded to more accurate recognition.[3] Suggestibility was the number of misinformation items misattributed to the event, measured by the sum of misinformation misattributed to the both and event-only options (range = 0–5).[4] Finally, source monitoring was the sum of witnessed items (i.e., experienced in both, event only, or misinformation only) attributed to their correct source, as a proportion of total hits (range = 0–1; see Foley & Johnson, 1985). Descriptive statistics for these measures are displayed in Table 1, and a further table of attribution patterns for all four sources is provided in Appendix B in the supplemental material. Recognition performance in the nonmisinformed group was calculated as the sum of misses and false alarms.

[3] For the misinformed group, hits were equal to the sum of both, film, and story items attributed to any one of the both, film, or story options at test. Misses were equal to the sum of both, film, and story items attributed to the new option at test. False alarms were equal to the sum of new items attributed to any one of the both, film, or story options at test.

[4] We were interested in the magnitude of the suggestibility effect (i.e., the quantity of items that were experienced in the misinformation phase that were misattributed to the event phase). However, it should be noted that in Studies 1 and 2, misattributions of misinformation details to the event significantly exceeded misattribution of new details to the event. Hence, in both studies, the suggestibility effect was significant.

Theory-of-Mind Performance

A latent trait analysis (with LAMI software, Bartholomew et al., 2002) was performed to examine whether Welch-Ross's (2000) two-component conceptualization of theory-of-mind development was statistically justified or whether all tasks shared a common underlying single component. A one-factor model was initially generated.

The proportion of the sample passing each task, together with standardized factor loadings for a single-factor solution are presented in Table 2. The pass rates revealed that the pretend–real task was passed less often than the false-belief task (see also individual response patterns in Appendix C in the supplemental material), even though we expected that the pretend–real task would be easier than the false-belief task, given the previous evidence (Flavell et al., 1987). The factor loadings were positive and fairly high for the majority of individual theory-of-mind measures and, in particular, for inference neglect and origins of belief. Goodness-of-fit was assessed by inspecting two-way margins.[5] This inspection indicated that a one-factor model provided an adequate fit for these data, accounting for 56% of the variance in individual task performance ($G^2 = 55.96$), and an additional second factor contributed only a further 9% to the proportion of explained G^2. This finding suggests commonality across tasks and justified computing a composite score based on summation of performance across all tasks rather than dividing into dual components. Because the latent trait analysis provided information about the relative degree of loading of individual tasks onto this factor, performance patterns were weighted accordingly. Hence, the composite score for each individual was not simply based on the sum of the binary scores for all tasks (0–6) but on a single component score, which took into account the relative contribution of each task. Performance patterns across tasks and corresponding component scores based on weighted summation of individual task performance are provided in Appendix C in the supplemental material.

Even though the two-factor solution was rejected, we nevertheless forced a two-factor model and inspected the factor loadings to assess Welch-Ross's (2000) dual conceptualization of theory of mind. The factor loadings for a two-factor solution are displayed in Table 3. Only one of the individual tasks (aspectuality) loaded more onto this second factor, and the loading value for all other tasks was relatively small after varimax and promax rotation. The pattern of loadings was not consistent with the two-component approach (conflicting mental representations vs. origins of knowledge) expected by Welch-Ross's model. In other words, each of these components did not load high on one factor and low on the other.

Table 2
Latent Trait Analyses: Pass Rates and Standardized Factor Loadings for Single Factor Solutions

	Study 1		Study 2	
Theory-of-mind task	Pass rate	Standardized factor loading	Pass rate	Standardized factor loading
Pretend–real	.63	.69	.87	.96
False belief (other)	.73	.76	.64	.96
False belief (own)	.65	.83	.60	.87
Origins of belief	.65	.89	.67	.99
Aspectuality	.35	.86	.31	.95
Inference neglect	.24	.90	.24	.99

Table 3
Latent Trait Analyses: Standardized Factor Loadings for Two-Factor Solutions

	Study 1 factor loadings		Study 2 factor loadings	
Theory-of-mind task	Factor 1	Factor 2	Factor 1	Factor 2
Pretend–real[a]	.78	−.21	.62	.77
False belief (other)[a]	.78	−.06	.92	.30
False belief (own)[a]	.88	−.10	.98	−.18
Origins of belief[b]	.92	.02	.83	.55
Aspectuality[b]	.50	.86	.88	.35
Inference neglect[b]	.72	.56	.92	.37

[a]Corresponds to Welch-Ross's (2000) conflicting mental representations component. [b]Corresponds to Welch-Ross's (2000) origins-of-knowledge component.

Associations Between Memory and Theory of Mind

Misinformed group. Pearson's *r* correlations between dependent measures for the misinformed group are displayed in Table 4 (top half of matrix). Recognition and source monitoring shared only a weak association, implying that these measures are partially independent (see Johnson et al., 1993). Theory of mind and suggestibility were strongly related to chronological and verbal mental age, in addition to all three measures of memory performance, highlighting the importance of accounting for both chronological age and verbal mental age when examining the association between theory of mind and suggestibility.

A series of nested multiple regression analyses were used to establish whether theory of mind explained unique variation in memory performance beyond that accounted for by chronological and verbal mental age.[6] A separate regression analysis was conducted for recognition, suggestibility, and source monitoring (see Table 5). Each of the three predictors was entered into the model in a separate block, and the order of entry was identical for all regressions: chronological age, verbal mental age, theory of mind.

Chronological and verbal mental age combined accounted for approximately 30% of the variance in recognition, and theory of mind contributed only a further 1% to the model. Similarly, theory of mind was not reliably associated with source monitoring and

[5] With these particular data, a large number of expected frequencies had values less than five, and it is recommended in this case that we look at the fit of *two-way margins* (see Bartholomew et al., 2002, pp. 184–186). Judgments of best fit are based on the chi-square residuals for each of the associations between two variables at a time.

[6] Although it is often ill advised to use regression analyses with such high levels of collinearity, in this particular case it is warranted, because we were interested in whether theory of mind explained variance in memory performance beyond that accounted for by both age measures. Here, we assumed that theory of mind measures a developmental progression that shares some qualities with verbal mental ability but nevertheless measures a separate aspect of development.

Table 4
Pearson's r Correlations for Study 1 (Misinformed Group) and Study 2

Measure	Recognition	Suggestibility	Source monitoring	Chronological age	Verbal mental age	Theory of mind
Recognition	—	.11	−.24	−.54**	−.46**	−.35**
Suggestibility	.01	—	−.80**	−.43**	−.53**	−.59**
Source monitoring	−.22	−.71**	—	.50**	−.60**	.60**
Chronological age	−.36*	−.41**	.47**	—	.82**	.76**
Verbal mental age	−.32*	−.47**	.45**	.79**	—	.78**
Theory of mind	−.42**	−.59**	.66**	.73**	.75**	—

Note. Study 1 (misinformed group) data are displayed above the diagonal, and Study 2 data are displayed below the diagonal.
*$p < .05$. **$p < .01$ (two-tailed); $n = 36$ (Study 1), $n = 45$ (Study 2).

contributed only a nonsignificant 5% to the variance after accounting for both age measures. Hence, even though theory of mind was correlated with recognition and source monitoring (see Table 4), the associations were driven largely by variations in verbal ability. In contrast, theory of mind accounted for a significant 10% of the variance in suggestibility, once chronological and verbal mental age were partialed out, thus indicating that theory of mind did indeed share unique variance with suggestibility.

Taken together with the suggestibility regression analysis, the source-monitoring findings imply that if theory of mind does mediate source monitoring in some way, it may be specific to misattributions of items to a previously experienced source, measured in this case by suggestibility. However, it is possible that the lack of a unique relationship between source monitoring and theory of mind may be an artifact of the intervals imposed between the three stages of the eyewitness procedure. Because the misinformation-to-test interval (10 min) was much smaller than the event-to-misinformation delay (1 day), misattributions to the recently experienced misinformation narrative may have been unlikely, relative to misattributions to the event seen the day before. This may have created an artificial imbalance between the suggestibility and source-monitoring measures, because only the latter included attributions in both directions (i.e., from the event to the misinformation and vice versa). As a consequence, we inspected source attributions further with a paired samples *t* test to examine whether misattributing a previously seen item to the new option was dependent on where it was originally witnessed (event only vs. misinformation only; see Appendix B in the supplemental material for means). Event-only items were indeed significantly more likely than misinformation-only items to be labeled as *new,* $t(35) = 3.61$, $p = .001$, indicating that children did indeed appear to forget details from the event more readily than those from the more recently experienced misinformation phase. This forgetting bias and the potential impact that it may have had on the relationship between our measures of source monitoring and theory of mind were addressed in Study 2.

Table 5
Multiple Regression Analyses: Misinformed Group (Study 1)

Measure and predictor entered	R^2	R^2 change	F	df	p
Recognition					
1. Chronological age	.295	.295	14.25	1, 34	.001
2. Verbal mental age	.296	>.001	0.01	1, 33	.922
3. Theory of mind	.309	.013	0.61	1, 32	.442
Suggestibility					
1. Chronological age	.183	.183	7.61	1, 34	.009
2. Verbal mental age	.281	.099	4.53	1, 33	.041
3. Theory of mind	.383	.101	5.25	1, 32	.029
Source monitoring					
1. Chronological age	.252	.252	11.46	1, 34	.002
2. Verbal mental age	.355	.103	5.25	1, 33	.029
3. Theory of mind	.403	.048	2.58	1, 32	.118

Not misinformed group. The relationship between theory of mind and recognition ($M = 2.58$, $SD = 2.02$, range = 0–8) was also examined for the not misinformed group. Pearson's correlations between recognition and chronological age ($r = -.51$), verbal mental age ($r = -.61$), and theory of mind ($r = -.50$) for the not misinformed group were strong and significant. A series of nested multiple regression analyses showed that theory of mind ($\Delta R^2 = .04$, $p = .18$) did not make a significant contribution to variation in recognition beyond that accounted for by chronological ($R^2 = .04$, $p < .1$) and verbal mental age ($\Delta R^2 = .12$, $p < .05$). This mirrors the recognition analysis findings for the misinformed group, illustrating that theory of mind is not associated with recognition, whether the participant is misinformed or not.

Discussion

Study 1 built on previous research (Welch-Ross, 1999, 2000; Welch-Ross et al., 1997) and established that both conflicting mental representation and origins-of-knowledge components of theory of mind can be legitimately summed within a single composite score. In addition, only suggestibility (but not recognition or source monitoring) shared a unique association with theory of mind, after accounting for both chronological and verbal mental age. Although these findings support previous research indicating a link between theory of mind and suggestibility, they contradict the mental-state reasoning model (Welch-Ross, 2000) in two ways. First, the findings are contrary to Welch-Ross's (1999, 2000; Welch-Ross et al., 1997) dichotomization of theory-of-mind performance. Second, the findings do not support Welch-Ross's speculation about the specific role of source monitoring in mediating associations between theory of mind and suggestibility. However, two issues needed to be addressed in a further study. First, the pretend–real task did not pick up skills prior to false-belief understanding, which may have reduced the sensitivity of the theory-of-mind measure. Second, a forgetting bias may have masked an association between source monitoring and theory of mind.

Study 2

The purpose of Study 2 was to reduce the forgetting biases seen in Study 1 and to use a more sensitive theory-of-mind measure, to ensure that unique associations between theory of mind and suggestibility (but not source monitoring or recognition) were not artifacts of the procedure. Study 1 was repeated with a new sample, with two modifications.

To rectify the forgetting bias, we wanted to ensure that both event and misinformation phases were equally likely to be forgotten at test. However, this is not easily achievable given that the misinformation always follows the event and is, therefore, always more recent at the point of retrieval. One way around this is to apply the notion of the ratio rule outlined in temporal distinctiveness theories of memory (Neath & Crowder, 1990). According to these accounts, when items are presented further away in time from test, they become less temporally distinct. In other words, the difference in presentation time between the event and the misinformation would appear to diminish if both these phases were presented close together and further away from test, thus reducing the magnitude of the recency effect (Nairne, Neath, Serra, & Byun, 1997). Moreover, the apparent temporal distance between items at test is quantifiable and is measured by the ratio of the interitem presentation interval and the retention interval between last item and test (Brown, Neath, & Chater, 2007). In the present study, we extended the size of the misinformation–test interval relative to the event–misinformation interval to reduce the recency effect. We applied this same ratio over two event–test time periods.

The second modification was to the pretend–real task. In Study , the pretend–real task did not distinguish skills required prior to an understanding of false belief, as we might expect from previous findings (Flavell et al., 1987). This may be attributed to the particular procedure we adopted in Study 1, which measured the ability to distinguish pretend and real objects (Welch-Ross et al., 1997). This task is qualitatively different from that of Flavell et al. (1987), who asked children to use object substitution in situations of pretence. This vital difference may have contributed to the unexpectedly low pass rate. In view of this, the pretend–real task was modified to follow more closely the procedure outlined by Flavell et al., described later.

Method

Participants

Forty-five participants (17 male, 28 female) between ages 3 and 7 years ($M = 64.16$ months, $SD = 14.50$, range = 40–89 months) were drawn from a state school and nursery. In line with the first study, the school was selected to ensure that participants were recruited from areas of average socioeconomic status, but details of parents' education, occupation, and income were not recorded. English was the native language of all participants. There were similar numbers of children between 3.0 and 4.5 years old ($n = 13$), between 4.5 and 6.0 years old ($n = 19$), and between 6.0 and 7.5 years old ($n = 13$). Children were allocated to short and long event-test groups on the basis of maintaining a similar range of chronological age across groups (short group: $M = 63.16$ months, $SD = 14.44$, range = 40–89 months; long group: $M = 65.40$ months, $SD = 14.86$, range = 40 –89 months) and verbal mental age (short group: $M = 57.88$ months, $SD = 20.13$, range = 25–94 months; long group: $M = 59.30$ months, $SD = 15.09$, range = 25–9 months).

Materials and Procedure

The procedure was identical to that of Study 1, except for the following modifications. First, all children watched the same version of the film because previous counterbalancing indicated no effect of item type.

Second, the event–misinformation interval was small relative to the misinformation–test interval, with an event–misinformation to misinformation–test interval ratio of approximately 1:6. In addition, children were divided into short and long event–test delay conditions. In the short event–test group, the event and misinformation were experienced on the same day (average event–misinformation interval = 7 hr), and testing was carried out 2 days later (average misinformation–test interval = 42 hr). In the long event–test delay group, the event and misinformation stages were separated by 1 day (event–misinformation interval = 1 day), and children were tested 6 days later (misinformation–test = 6 days). Hence, the delay ratio (1:6) remained constant across the two groups, despite differences in event–test delays.

The third modification was the pretend–real procedure. The child was shown an object (e.g., a sponge) and was asked the reality test question, "What is this really and truly?" The experimenter then explained that she was going to pretend that the object was something else (e.g., a truck) and then performed an appropriate action. The experimenter then posed the pretense test question, "What am I pretending this is right now?" Children completed two practice trials with feedback from the experimenter (a sponge and a sheet of paper representing a truck and a blanket, respectively) and four test trials (brush, toothbrush, cup, and comb representing a snake, telephone, hat, and boat, respectively). Children passed each trial by answering both reality and pretense questions correctly. An overall pass was obtained if all test trials were completed successfully.

Results

Eyewitness Memory Performance

Independent sample t tests examined whether memory performance differed significantly across the short and long event–test delay group (see Table 1 for means). Recognition, $t(43) = -0.81$, $p = .42$, suggestibility, $t(43) = 1.64$, $p = .11$, and source discrimination, $t(43) = -0.59$, $p = .56$, did not vary significantly across the two groups. Thus, data were collapsed across the groups in further analysis. An attribution table is provided in Appendix B in the supplemental material.

A paired sample t test examined whether the forgetting bias identified in Study 1 was addressed by modifications to the ratio of time intervals. In contrast to Study 1, event-only items ($M = 0.67$, $SD = 0.71$) were not significantly more likely than misinformation-only items ($M = 0.67$, $SD = 0.71$) to be attributed to the new option at test, $t(44) = 0.16$, $p = .88$. Hence, by placing the event and misinformation at a similar temporal distance from test, from the perspective of the retrieval phase, forgetting biases were overcome.

Theory-of-Mind Performance

A latent trait analysis assessed whether individual theory-of-mind task performance was explained by a single latent factor. The pass rate and factor loadings for a one-factor solution are displayed in Table 2. In contrast to Study 1, the modified pretend–real task had a higher pass rate than the false-belief tasks. Indeed, the overall pattern across tasks indicated that as the expected difficulty of tasks was progressively increased, the pass rate tended to decrease. Moreover, the individual response patterns (see Appendix C in the supplemental material) were also much more consistent with expected order of success than in Study 1. Table 2 also shows that the factor loadings were all positive and high for a one-factor solution, as confirmed by a two-way margin goodness-of-fit analysis that showed that this single latent factor explained a large proportion of the variance of individual theory-of-mind performance (G^2 explained = 87%). A second factor did not add significantly to the model (5% of G^2 explained). In line with Study 1, we examined the pattern of loading for a two-factor solution (Table 3). After varimax and promax rotation, the loadings on the second factor were all low except for pretend–real performance. Critically, and in line with Study 1, the loading patterns did not fit with Welch-Ross's (2000) dual-component conceptualization of theory of mind and did not suggest a separation of conflicting mental representation and origins-of-knowledge components.

Associations Between Memory and Theory of Mind

Pearson's *r* correlations among memory, theory-of-mind, and age measures are displayed in Table 4 (bottom half of the matrix). Recognition and source monitoring were weakly related and were thus partially independent. Chronological and verbal mental age and theory of mind were highly related to all three memory measures and, in particular, to source monitoring.

Nested multiple regression analyses (see Study 1) were conducted to examine whether theory of mind predicted memory performance, independent of chronological and verbal mental age (see Table 6). Theory of mind shared unique variance with suggestibility (13% explained variance) but not with recognition after partialing out age factors. However, in contrast to Study 1, source monitoring was also uniquely associated with theory of mind, accounting for around 21% of the variance after controlling for chronological and verbal mental age. Hence, once forgetting biases were reduced, theory of mind was not only uniquely related to suggestibility but also to source monitoring, implying that that source monitoring, but not old–new recognition, may be implicated specifically in the association between theory of mind and suggestibility.

In a supplementary analysis, we divided summed theory-of-mind performance into the two separate components (conflicting mental representation and origins of knowledge) described by Welch-Ross (2000). The magnitudes of nonparametric correlations between theory-of-mind and source-monitoring performance were similar for both origins-of-knowledge (spearman's $\rho = .55$) and conflicting-mental-representation performance (spearman's $\rho = .49$) components of theory of performance. This provides further support that the relationship between source monitoring and theory of mind was not specific to so-called origins-of-knowledge performance but also extended to conflicting-mental-representation performance.

In a final regression analysis, suggestibility and source monitoring were entered first into the model as separate blocks, followed by theory of mind (see Table 7). Suggestibility accounted for 35% of the variance in theory of mind, but source monitoring added a significant further 11% to the model. A reversal of entry of predictors demonstrated that after accounting for the variance shared with source monitoring (43%), suggestibility explained only a further nonsignificant 3%. In other words, the relationship between theory of mind and source monitoring was not dependent on the degree of suggestibility, but the association between theory of mind and suggestibility was underpinned by concurrent improvements in source monitoring.[7]

Table 6
Multiple Regression Analyses (Study 2)

Measure and predictor entered	R^2	R^2 change	*F*	*df*	*p*
Recognition					
1. Chronological age	.130	.130	6.54	1, 43	.015
2. Verbal mental age	.134	.004	0.18	1, 42	.672
3. Theory of mind	.183	.048	2.43	1, 41	.127
Suggestibility					
1. Chronological age	.169	.169	8.74	1, 43	.005
2. Verbal mental age	.225	.056	3.02	1, 42	.090
3. Theory of mind	.355	.131	8.32	1, 41	.006
Source monitoring					
1. Chronological age	.216	.216	11.87	1, 43	.001
2. Verbal mental age	.232	.016	0.85	1, 42	.361
3. Theory of mind	.440	.208	15.19	1, 41	.001

General Discussion

The current research had three aims. The first was to test Welch-Ross's (2000) dual-component approach to theory-of-mind development with latent trait analysis. The second aim was to identify the memory processes that underpinned theory-of-mind and suggestibility associations. In particular, we explored whether the association was mediated by improvements in source monitoring, as suggested in Welch-Ross's (2000) mental-state reasoning

[7] One reviewer suggested that the source-memory and origins-of-belief tasks were highly similar and that this may have driven the association between theory of mind and source memory. However, there was a vital difference, albeit a subtle one, between the tasks. The source-memory task measured the child's memory for the context of experienced events, whereas the origins-of-belief task was designed to tap the understanding that mental states (i.e., current beliefs) are linked to and are dependent on perceptions. Hence, the memory task required the child to identify the context of the memory (*where* did you see *x*?) and simply measured memory performance, whereas the mental-state task required the child to identify the causal link between mental states and perceptions (*how* do you know *x*?) and was designed to measure conceptual, mental-state understanding. Having said this, as the reviewer pointed out, this difference is not readily apparent within the literature. It is important to note that the theory-of-mind composite score reported here was not driven solely or, indeed, biased by origins-of-belief performance. The factor loading for origins of belief was fairly similar to the loading for performance on other theory-of-mind tasks, thus indicating that origins-of-belief performance did not contribute excessively to the composite score.

Table 7
Multiple Regression Analyses to Investigate Whether Source Monitoring Underpinned the Association Between Theory of Mind and Suggestibility (Study 2)

Model and variable entered	R^2	R^2 change	F	df	p
Model 1					
1. Suggestibility	.351	.351	23.24	1, 43	.001
2. Source monitoring	.464	.113	8.83	1, 42	.005
Model 2					
1. Source monitoring	.433	.433	32.82	1, 43	.001
2. Suggestibility	.464	.031	2.40	1, 42	.129

model, or was more generally mediated by old–new recognition skills. The third aim was to establish whether the association between improvements in theory-of-mind understanding and reductions in suggestibility across the preschool years was simply a reflection of improvements in verbal ability.

Dual or Single Component of Theory of Mind

Latent trait analyses indicated that a significant proportion of the variation in performance on six standard theory-of-mind tasks was explained by a single latent factor, and this finding was consistent across Studies 1 and 2. The proportion of variance explained by this single latent factor was particularly high following a modification to the pretend–real task, which was introduced in the final study. One possible explanation for the increased variance explained is that children in Study 2 tended to pass subsequent theory-of-mind tasks only when pretend–real skills were in place, resulting in a reduced number of response patterns, which may have been influenced the degree to which the latent factor explained manifest variables.

The single-factor finding is informative in two ways. First, a latent factor suggests a degree of commonality across the six separate theory-of-mind measures. Naturally, we are not suggesting that each of the individual theory-of-mind tasks measures an identical aspect of development, and clearly there are differences between the types of performance measured across tasks. Nevertheless, the single factor does suggest that there is an underlying aspect of development that is common to all of these particular measures of theory-of-mind performance. This finding fits well with Wellman and Liu's (2004) scaling of several theory-of-mind measures and with their suggestion that "theory-of-mind understandings represent an extended and progressive set of conceptual acquisitions" (p. 537) that form a consistent or single scale.

Second, for the purposes of our research, the single latent factor provides justification for combining performance on all six tasks within a single composite theory-of-mind measure. This was important to establish because Welch-Ross (2000) argued for a dual-component approach while combining all measures together in a single score, and we wanted to ascertain which of these was a valid approach. Indeed, the patterns of factor loading for an enforced two-factor solution did not correspond with those expected by a division into conflicting-mental-representation and origins-of-knowledge components, nor was there a substantial difference between the magnitude of the relationship between source monitoring and either of these components. Hence, the present data are not supportive of Welch-Ross's (2000) assertion that two components of theory-of-mind development (the ability to reason about conflicting mental representations and appreciation of the origins of knowledge) mediate suggestibility in specific ways.

Explaining the Nature of the Single Theory-of-Mind Factor

There are two reasons why we might have expected both conflicting-mental-representations and origins-of-belief components to load onto a single factor. First, as noted in the introduction, the division of theory-of-mind performance into these separate components is not clear within the theory-of-mind literature, and on the contrary, the two skills are assumed to be inherently linked. Second, false-belief understanding is likely to measure more than the ability to reason about conflicting mental representation. Although early accounts of false belief suggest that an ability to handle conflicting mental representations predicted success (Wimmer & Perner, 1983), Perner (1991) later suggested that it was the ability to model representations as representations that was the critical factor in passing this task. He proposed that handling multiple conflicting representations refers to a secondary understanding of representation, which allows children to make distinctions between pretend and real and to consider hypotheticals. This type of understanding is said to predate false-belief success, when meta-representational understanding is acquired.

Although the latent factor suggests a universal aspect within all theory-of-mind tasks, it is not informative about the exact nature of this shared variance. One possibility is that success on both false-belief and origins-of-knowledge tasks relies on the common manifestation of an understanding of the links between knowledge states and experience. Indeed, Wimmer and Gschaider (2000) suggested that because the concept of belief entails a causal component, the concept of belief is unable to mature until conceptual understanding of informational causation is developed. Hence, they argued that the understanding of belief is driven by an awareness of the causal connection between information and knowledge. In contrast, Perner (1991, 2000) suggested that meta-representational skills underpin both false-belief and origins-of-knowledge development. Indeed, Perner (2000; Perner & Ruffman, 1995) also explained the development of the ability to form episodic memories (which necessarily include information about source) in terms of meta-representation.

An alternative explanation is that the latent factor is tapping skills associated with theory-of-mind understanding, rather than theory of mind per se. Success on false-belief tasks has been linked to improvements in inhibitory control (Carlson & Moses, 2001) as well as to broader deficits in counterfactual reasoning (Riggs, Peterson, Robinson, & Mitchell, 1998) and embedded causal reasoning (Frye, Zelazo, & Palfai, 1995), and it is possible that these other skills may contribute to or indeed underpin associations across different aspects of theory-of-mind development. We return to this point at the end of the discussion.

The Role of Memory Processes and Verbal Ability

In the introductory section, we suggested that previous demonstrations of associations between theory of mind and suggestibility

(Welch-Ross, 2000) may have been driven by source monitoring or by improvements in old–new recognition performance. We also speculated that verbal ability may have played a significant role in driving apparent associations between theory of mind and memory measures. We discuss the findings of these two investigations together because they are necessarily linked. The findings were informative in several respects.

It is important to note that suggestibility shared unique variance with theory of mind, even after accounting for chronological and verbal mental age across two studies. This is a critical demonstration that the association between theory of mind and suggestibility is legitimate and is not a reflection of concurrent improvement in verbal ability.

We also explored whether theory of mind was associated specifically with source monitoring or with the less demanding skill of old–new recognition. In Study 1, theory of mind was not related to source monitoring after accounting for chronological or verbal mental age, but an inspection of source-attribution patterns suggested that this may have been due to a forgetting bias induced by the relative length of the time intervals used in the three-stage procedure. We addressed this in Study 2 by using a long misinformation-to-test interval relative to the event-to-misinformation delay, on the basis of the rationale provided by temporal distinctiveness models of forgetting (Brown et al., 2007). Once forgetting biases were minimized, both suggestibility and source monitoring shared a strong relationship with theory-of-mind performance independent of chronological and verbal mental age (Study 2). A final regression confirmed that the association between theory of mind and suggestibility was indeed underpinned by improvements in source monitoring.

We wanted to clarify whether this relationship also extended to recognition. Although recognition was correlated with theory of mind (Studies 1 and 2), the association was driven by changes in development with chronological and verbal mental age. This lack of a unique relationship between theory of mind and recognition also extended to circumstances where children were not misinformed (Study 1).

Taken together, these findings provide compelling evidence that the relationship between theory of mind and suggestibility is driven specifically by improvements in source monitoring and not by the development of old–new recognition skills. Moreover, source monitoring was not moderated solely by the development of origins-of-knowledge skills, contrary to Welch-Ross's (2000) proposal. Instead, overall theory-of-mind development (i.e., both conflicting-mental-representation and origins-of-knowledge performance) contributed to improvements in source monitoring.

Our findings are consistent with previous research linking source-memory accuracy with false-belief success (Naito, 2003) and with work demonstrating that children with autism, who demonstrate deficits in theory-of-mind development, have poor source memory (Russell & Jarrold, 1999). A tentative conclusion is that theory-of-mind development encourages the ability to effectively monitor memory for source, which in turn leads to reductions in suggestibility across the preschool years. Of course, this direction of influence is speculative, because these particular data do not show whether theory of mind is a precursor for source monitoring or vice versa, or whether these two aspects of development appear concurrently. A longitudinal approach may help in addressing causality. Nevertheless, our findings do demonstrate a strong and specific association between theory of mind and source monitoring.

Explaining the Association Between Theory of Mind and Source Monitoring

Naturally, the question of why theory-of-mind development and memory for source are linked remains. We have contemplated several possibilities. First, both theory of mind and source monitoring may require specific causal links to be made between knowledge and the circumstances in which knowledge was acquired. Second, both require effortful reflection on one's own mental contents and rely on drawing conclusions about nonobservable relations through inference. For example, false-belief understanding relies on an understanding that mental states are not directly observable but are inferred from behavior. In a similar vein, source monitoring requires links to made between the contents of memory and the context in which they are acquired, but these relations are not directly observable and have to be inferred from integral features of the memory representation following reflection by means of an attribution process at retrieval (Johnson et al., 1993). Third, not only false-belief understanding, but also origins-of-knowledge task success and accurate source monitoring, may rely on meta-representational understanding. Perner (2000) suggested that the experiential awareness that accompanies episodic memories can be activated only when children are able to remember and represent events as personally experienced, and this requires meta-representational understanding. Indeed, it has been suggested that the aspectuality task requires the understanding that "knowledge can represent an object *as being a certain way*" (O'Neill et al., 1992, p. 488). Moreover, Mitchell and Johnson (2000) suggested that judgments about the source of a memory are based on "a representation of their constructed understanding of an event rather than a representation of the event itself" (p. 183). This observation sits well with Perner's formulation and suggests that it is not merely the ability to represent multiple models simultaneously that links theory of mind with suggestibility and source monitoring, but the ability to construct, evaluate, and reflect on memories as representations.

Having established a strong association between theory of mind and source monitoring, we add two caveats to our findings. It should be noted that children with poorer theory-of-mind understanding were not completely inept at monitoring the source of their memories. Although their performance on the source-memory task was significantly inferior to performance of those children with a good understanding of theory of mind, source-memory performance was nevertheless above chance level. Hence, identifying the source of their memories was clearly challenging, but not impossible, for children with poor theory-of-mind understanding. This suggests that theory of mind may help to moderate improvements in source memory but that it is not the only requirement for successful source attributions. Moreover, even though we controlled for verbal ability, it is still possible that some other factor associated with both source monitoring and theory of mind may be driving the relationship, and this may be worth considering in future research. In particular, we suggest that a fruitful area for exploration is the possible role of reasoning abilities and executive function. We expand on this next.

Executive Function and Links to Theory of Mind and Source Monitoring

A growing body of evidence suggests that theory-of-mind development may not merely be driven by conceptual change, but by the emergence of executive function (including inhibitory control, planning, and working memory) and reasoning skills, which are said to play a crucial role in both the acquisition and the expression of the concepts necessary to pass theory-of-mind tasks.

For example, Frye, Zelazo, Brooks, and Samuels (1996) suggested that false-belief success is, in part, reliant on the use of embedded rules in causal reasoning, and they reported that false-belief success was associated with the ability to switch between rules of sorting cards according to color or shape (Frye et al., 1995). Because source monitoring requires deliberation about several dimensions of information contained within the memory representation, it is conceivable that an inability to use embedded rules may influence the efficacy of source judgment, particularly if the child is not able to consider information and make inferences along several dimensions simultaneously.

In addition, inhibitory control (the ability to disengage from a salient or prepotent response) has been strongly linked to theory-of-mind performance (Carlson & Moses, 2001; Carlson, Moses, & Claxton, 2004; Hughes, 1998). Moreover, source-monitoring accuracy (Roberts & Powell, 2005; Ruffman, Rustin, Garnham, & Parkin, 2001; Scullin & Bonner, 2006) is also associated with improved performance on inhibitory control tasks. One explanation is that accurate source monitoring relies on the inhibition of decisions based on familiarity at retrieval (Roberts & Powell, 2005). It therefore seems plausible that the relationship between source monitoring and theory of mind reported in the current article may be driven by this third factor. Nevertheless, a recent study that examined the relationship between inhibitory control and suggestibility indicated that the relationship was weak once language ability was controlled for (Roebers & Schneider, 2005). Hence, we cannot assume that the associations among theory of mind, source monitoring, and suggestibility are necessarily driven by inhibitory control, but an exploration of the role of inhibitory control in this three-way relationship is strongly recommended for future research.

Having noted some possible areas of future investigation, we emphasize that what the current data do demonstrate is a substantial association between theory of mind and suggestibility, even after accounting for improvements in verbal ability. Moreover, the finding that this association is linked specifically with source monitoring and not recognition adds significantly to our understanding of the memory processes driving the relationship.

References

Ackil, J. K., & Zaragoza, M. S. (1995). Developmental differences in eyewitness suggestibility and memory for source. *Journal of Experimental Child Psychology, 60,* 57–83.

Astington, J. W. (1993). *The child's discovery of the mind.* Cambridge, MA: Harvard University Press.

Astington, J. W., & Gopnik, A. (1988). Knowing you've changed your mind: Children's understanding of representational change. In A. W. Astington, P. L. Harris, & D. R. Olson (Eds.), *Developing theories of mind* (pp. 193–206). Cambridge, England: Cambridge University Press.

Ayers, M. S., & Reder, L. M. (1998). A theoretical review of the misinformation effect: Predictions from an activation-based memory model. *Psychonomic Bulletin & Review, 5,* 1–21.

Baron-Cohen, S., Leslie, A. M., & Frith, U. (1985). Does the autistic child have a theory of mind? *Cognition, 21,* 37–46.

Bartholomew, D. J., Steele, F., Moustaki, I., & Galbraith, J. I. (2002). *The analysis and interpretation of multivariate data for social scientists.* Boca Raton, FL: CRC Press. LAMI software retrieved June 2, 2008, from http://www.cmm.bris.ac.uk/team/aimdss.shtml

Bright-Paul, A., Jarrold, C., & Wright, D. B. (2005). Age-appropriate cues facilitate source-monitoring and reduce suggestibility in 3- to 7-year-olds. *Cognitive Development, 20,* 1–18.

Brown, G. D. A., Neath, I., & Chater, N. (2007). A temporal ratio model of memory. *Psychological Review, 114,* 539–576.

Bruck, M., & Ceci, S. J. (1999). The suggestibility of children's memory. *Annual Review of Psychology, 50,* 419–439.

Bruck, M., & Melnyk, L. (2004). Individual differences in children's suggestibility: A review and synthesis. *Applied Cognitive Psychology, 18,* 947–996.

Carlson, S. M., & Moses, L. J. (2001). Individual differences in inhibitory control and children's theory of mind. *Child Development, 72,* 1032–1053.

Carlson, S. M., Moses, L. J., & Claxton, L. J. (2004). Individual differences in executive functioning and theory of mind: An investigation of inhibitory control and planning ability. *Journal of Experimental Child Psychology, 87,* 299–319.

Ceci, S. J., & Bruck, M. (1993). Suggestibility of the child witness—An historical review and synthesis. *Psychological Bulletin, 113,* 403–439.

Dennett, D. C. (1978). Beliefs about beliefs. *Behavioural and Brain Sciences, 4,* 568–570.

Dunn, L. M., Dunn, L. M., & Whetton, C. (1997). *British picture vocabulary scale* (2nd ed.). Windsor, England: NFER-Neslon.

Farrant, A., Blades, M., & Boucher, J. (1998). Source monitoring by children with autism. *Journal of Autism and Developmental Disorders, 28,* 43–50.

Flavell, J. H., Flavell, E. R., & Green, F. L. (1987). Young children's knowledge about the apparent–real and pretend–real distinctions. *Developmental Psychology, 23,* 816–822.

Foley, M. A., & Johnson, M. K. (1985). Confusions between memories for performed and imagined actions: A developmental comparison. *Child Development, 56,* 1145–1155.

Frye, D., Zelazo, P. D., Brooks, P. J., & Samuels, M. C. (1996). Inference and action in early causal reasoning. *Developmental Psychology, 32,* 120–131.

Frye, D., Zelazo, P. D., & Palfai, T. (1995). Theory of mind and rule-based reasoning. *Cognitive Development, 10,* 483–527.

Giles, J. W., Gopnik, A., & Heyman, G. D. (2002). Source monitoring reduces the suggestibility of preschool children. *Psychological Science, 13,* 288–291.

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

Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. *Psychological Bulletin, 114,* 3–28.

Lindsay, D. S., & Johnson, M. K. (1989). The eyewitness suggestibility effect and memory for source. *Memory & Cognition, 17,* 349–358.

Loftus, E. F. (1979). *Eyewitness testimony.* Cambridge, MA: Harvard University Press.

Loftus, E. F., Miller, D. G., & Burns, H. J. (1978). Semantic integration of verbal information in visual memory. *Journal of Experimental Psychology: Human Learning and Memory, 4,* 19–31.

Mitchell, K. J., & Johnson, M. K. (2000). Source monitoring. Attributing mental experiences. In E. Tulving & F. I. M. Craik (Eds.), *The Oxford handbook of memory* (pp. 179–195). Oxford, England: Oxford University Press.

Mitchell, P., & Lacohee, H. (1991). Children's early understanding of false belief. *Cognition, 39,* 107–127.

Morrison, C. M., Chappell, T. D., & Ellis, A. W. (1997). Age of acquisition norms for a large set of object names and their relation to adult estimates and other variables. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 50*(A), 528–559.

Nairne, J. S., Neath, I., Serra, M., & Byun, E. (1997). Positional distinctiveness and the ratio rule in free recall. *Journal of Memory and Language, 37,* 155–166.

Naito, M. (2003). The relationship between theory of mind and episodic memory: Evidence for the development of autonoetic consciousness. *Journal of Experimental Child Psychology, 85,* 312–336.

Neath, I., & Crowder, R. G. (1990). Schedules of presentation and temporal distinctiveness in human memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16,* 316–327.

O'Neill, D. K., Astington, J. W., & Flavell, J. H. (1992). Young children's understanding of the role that sensory experiences play in knowledge acquisition. *Child Development, 63,* 474–490.

O'Neill, D. K., & Gopnik, A. (1991). Young children's ability to identify the sources of their beliefs. *Developmental Psychology, 27,* 390–397.

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

Perner, J. (2000). Memory and theory of mind. In E. Tulving & F. I. M. Craik (Eds.), *The Oxford handbook of memory* (pp. 297–312). Oxford, England: Oxford University Press.

Perner, J., & Ruffman, T. (1995). Episodic memory and autonoetic consciousness—Developmental evidence and a theory of childhood amnesia. *Journal of Experimental Child Psychology, 59,* 516–548.

Poole, D. A., & Lindsay, D. S. (1995). Interviewing preschoolers—Effects of nonsuggestive techniques, parental coaching, and leading questions on reports of nonexperienced events. *Journal of Experimental Child Psychology, 60,* 129–154.

Quas, J. A., Qin, J. J., Schaaf, J. M., & Goodman, G. S. (1997). Individual differences in children's and adults' suggestibility and false event memory. *Learning and Individual Differences, 9,* 359–390.

Quas, J. A., & Schaaf, J. M. (2002). Children's memories of experienced and nonexperienced events following repeated interviews. *Journal of Experimental Child Psychology, 83,* 304–338.

Riggs, K. J., Peterson, D. M., Robinson, E. J., & Mitchell, P. (1998). Are errors in false belief tasks symptomatic of a broader difficulty with counterfactuality? *Cognitive Development, 13,* 73–90.

Roberts, K. P., & Powell, M. B. (2005). The relation between inhibitory control and children's eyewitness memory. *Applied Cognitive Psychology, 19,* 1003–1018.

Roebers, C. M., & Schneider, W. (2005). Individual differences in young children's suggestibility: Relations to event memory, language abilities, working memory, and executive functioning. *Cognitive Development, 20,* 427–447.

Ruffman, T., Rustin, C., Garnham, W., & Parkin, A. J. (2001). Source monitoring and false memories in children: Relation to certainty and executive functioning. *Journal of Experimental Child Psychology, 80,* 95–111.

Russell, J., & Jarrold, C. (1999). Memory for actions in children with autism: Self versus other. *Cognitive Neuropsychiatry, 4,* 303–331.

Scullin, M. H., & Bonner, K. (2006). Theory of mind, inhibitory control, and preschool-age children's suggestibility in different interviewing contexts. *Journal of Experimental Child Psychology, 93,* 120–138.

Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory, 6,* 174–215.

Sodian, B., & Wimmer, H. (1987). Children's understanding of inference as a source of knowledge. *Child Development, 58,* 424–433.

Sparrevohn, R., & Howie, P. M. (1995). Theory of mind in children with autistic disorder: Evidence of developmental progression and the role of verbal ability. *Journal of Child Psychology and Psychiatry, 36,* 249–263.

Templeton, L. M., & Wilcox, S. A. (2000). A tale of two representations: The misinformation effect and children's developing theory of mind. *Child Development, 71,* 402–416.

Thierry, K. L., & Spence, M. J. (2002). Source-monitoring training facilitates preschoolers' eyewitness memory performance. *Developmental Psychology, 38,* 428–437.

Thomsen, Y., & Bernstein, D. (2005). Knowing that I didn't know: Preschoolers' understanding of their own false belief is a predictor of assents to fictitious events. *Applied Cognitive Psychology, 19,* 507–527.

Varouxaki, A., Freeman, N. H., Peters, D., & Lewis, C. (1999). Inference neglect and ignorance denial. *British Journal of Developmental Psychology, 17,* 483–499.

Welch-Ross, M. K. (1999). Preschoolers' understanding of mind: Implications for suggestibility. *Cognitive Development, 14,* 101–131.

Welch-Ross, M. K. (2000). A mental-state reasoning model of suggestibility and memory source monitoring. In K. P. Roberts & M. Blades (Eds.), *Children's source monitoring* (pp. 227–256). Mahwah, NJ: Erlbaum.

Welch-Ross, M. K., Diecidue, K., & Miller, S. A. (1997). Young children's understanding of conflicting mental representation predicts suggestibility. *Developmental Psychology, 33,* 43–53.

Wellman, H. M., Cross, D., & Watson, J. (2001). Meta-analysis of theory of mind development: The truth about false belief. *Child Development, 72,* 655–684.

Wellman, H. M., & Liu, D. (2004). Scaling of theory-of-mind tasks. *Child Development, 75,* 523–541.

Wimmer, H., & Gschaider, A. (2000). Children's understanding of belief: Why is it important to understand what happened? In P. Mitchell & K. J. Riggs (Eds.), *Children's reasoning and the mind* (pp. 253–266). Hove, England: Psychology Press.

Wimmer, H., Hogrefe, J., & Sodian, B. (1988). A second stage in children's conception of mental life: Understanding informational accesses as origins of knowledge and belief. In A. W. Astington, P. L. Harris, & D. R. Olson (Eds.), *Developing theories of mind* (pp. ??–??). Cambridge, England: Cambridge University Press.

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.

Yirmiya, N., Erel, O., Shaked, M., & Solomonica-Levi, D. (1998). Meta-analyses comparing theory of mind abilities of individuals with autism, individuals with mental retardation, and normally developing individuals. *Psychological Bulletin, 124,* 283–307.

Received December 13, 2006
Revision received January 17, 2008
Accepted January 23, 2008 ■