Journal of Experimental Child Psychology **78**, 107–129 (2001)
doi:10.1006/jecp.2000.2562, available online at http://www.idealibrary.com on IDEAL®

Video Reminders in a Representational Change Task: Memory for Cues but Not Beliefs or Statements

Philip David Zelazo and Janet J. Boseovski

University of Toronto, Toronto, Ontario, Canada

Two experiments investigated the effect of video reminders on 3-year-olds' performance in a representational change task. In Experiment 1, children in a video support condition viewed videotapes of their initial incorrect statements about a misleading container prior to being asked to report their initial belief. Children in a control condition viewed an irrelevant videotape. Despite reporting what they had said on the videotape, children in the video support condition typically failed the representational change task. Experiment 2 replicated the main findings from Experiment 1 and also revealed that a video reminder failed to increase the likelihood that children would correctly report what they had said about the object. Results are discussed in terms of the processes whereby mnemonic cues might affect performance on tasks assessing theory of mind.

Key Words: theory of mind; external representations; children's memory; interference; realist bias; executive function; video.

The recollection of false beliefs is one of the most widely assessed aspects of memory in preschoolers. In one task, called the representational change task (Gopnik & Astington, 1988), children are shown a familiar container (e.g., a Smarties box) and asked what it contains. Subsequently, the container is opened to reveal something unexpected (e.g., string) and children are asked to recall their initial expectation about its contents: "What did you think was in the box before I opened it?" Although most 4- and 5-year-olds answer the representational change question correctly, 3-year-olds often respond incorrectly, stating, for example, "String."

In order to answer the representational change question correctly, children must be able to recollect (or reconstruct) their initial false belief and inhibit any tendency that they may have to report their current belief. Failures to answer the question correctly could be attributed to limitations on episodic memory, autobiographical memory, source memory, working memory, inhibition, and/or rule use, among other things. However, because similar age-related improvements in

The research reported here was supported in part by a grant from NSERC of Canada to P. D. Zelazo. We thank L. Booth, A.-L. Cohen, D. Collister, N. Kirkham, S. Marcovitch, and M. Scharer for assisting with data collection. J. Burack, S. Jacques, and S. Marcovitch provided helpful comments on an initial draft of this article.

Address correspondence and reprint requests to P. D. Zelazo, Department of Psychology, University of Toronto, Toronto, Ontario, M5S 3G3, Canada. E-mail: zelazo@psych.utoronto.ca.

0022-0965/01 $35.00

performance have been observed in a variety of other tasks designed to assess children's ability to reason about mental states (e.g., Flavell, Flavell, & Green, 1983; Wimmer & Perner, 1983), recollection of false beliefs has often been interpreted in terms of the development of a folk psychology or a *theory of mind* (e.g., Gopnik, 1993; Wellman, 1990) rather than in terms of basic cognitive processes such as memory.

Recently, however, there has been considerable interest in identifying the cognitive processes that underlie children's performance on theory of mind tasks such as representational change. For example, several studies have documented correlations between theory of mind and various measures of executive function, including tests of rule use (Frye, Zelazo, & Palfai, 1995; Perner, Stummer, & Lang, 1999), response inhibition (Carlson, 1997), and working memory (Davis & Pratt, 1995; Gordon & Olson, 1998; Keenan, Olson, & Marini, 1998; see Perner & Lang, 1999, for a review). Although correlational, the results of this research are important because they locate theory of mind firmly within the broader context of profound changes in children's cognition that occur during the preschool period (e.g., see Bjorklund, 2000).

The problem of process modeling of performance on theory of mind tasks has also been explored experimentally. For example, several researchers have modified standard representational change tasks in an attempt to facilitate recollection of false beliefs or to reduce the salience of children's current (true) beliefs about reality (for reviews, see Freeman, 1994; Mitchell, 1994). Unfortunately, however, the results of this research have not been straightforward. Whereas some studies have demonstrated facilitation, others have not.

Mitchell and Lacohée (1991) reported an early success. In several experiments, children were shown a misleading container (e.g., a Smarties box containing pencils) and then asked to mail a picture that corresponded to their initial belief about the contents of the container (e.g., they posted a picture of Smarties into a toy postbox). In comparison to other children who mailed an irrelevant picture, these children were more likely to report their initial belief when reminded about the picture that they had mailed. The authors argued that the picture served as a salient, physical cue that helped children to remember, or at least to infer, their initial belief.

These findings were confirmed by Freeman and Lacohée (1995), who compared the efficacy of several different types of cue. In addition, however, these authors also found that the facilitative effect of mailing a picture was dependent on the use of a cued test question—a question that made explicit reference to the mailing event [e.g., "When you posted your picture in the postbox (touching closed postbox) what did you think was in here (holding forward closed Smartie tube)?"; Freeman & Lacohée, 1995, p. 42].

In contrast to these successes, however, Mitchell has twice failed to replicate the Mitchell and Lacohée (1991) results (see Freeman & Lacohée 1995, p. 55). Robinson, Riggs, and Samuel (1996) also failed to find support for the Mitchell and Lacohée (1991) results. These authors conducted a study using a similar

design except that instead of selecting and mailing a picture, children drew the supposed contents of a container. Although children correctly recalled the content of the cue, they nonetheless failed to acknowledge their initial incorrect belief. As Robinson and Mitchell (1994) noted, the literature on theory of mind is marked by numerous unexplained failures to replicate previous results, and it seems likely that some of the discrepancies can be attributed to unstable effects. However, it also seems likely that some of the discrepancies in the literature will be resolved once the boundary conditions of facilitation effects become clear.

At this point, it seems important to determine which cues are effective under which circumstances and for which types of test question. In the present experiments, we sought to determine whether we could improve 3-year-olds' performance on representational change tasks by providing them with a very strong form of mnemonic support: a videotape of their initial statement about a misleading object. The videotape is an especially strong cue because it is an undeniable physical record of their mistaken statement and because the inference from the recorded statement to the belief that motivated it (and especially to the initial statement itself) would seem, at least on the surface, to be more much direct than the inference from a picture that was mailed to an earlier mistaken belief. Although conceived independently, these experiments are similar to a study by Saltmarsh and Mitchell (1999), who found in two experiments that children who viewed a videotape of their initial incorrect statement performed better on a representational change question than children who did not. These results were obtained using both cued and uncued questions: specifically, in Experiment 2, Saltmarsh and Mitchell (1999) found evidence of facilitation even when they asked children a standard representational change question, "When I first showed you the box, before we opened it up, what did you think was inside?"

Despite the similarity between the current experiments and those of Saltmarsh and Mitchell (1999), there were several potentially important differences. First, we used uncued test questions and questions about reality in both experiments. The use of cued test questions raises the possibility that it is the cued question, rather than the event cue (e.g., the videotape), that accounts for any improvements in children's performance. Moreover, when children are asked cued test questions that refer to the event, it is possible that children are prompted to describe the event rather than their initial belief. The inclusion of the reality questions (used only in a between-subjects fashion in Saltmarsh & Mitchell, 1999, Experiment 1) was important because although most children fail theory of mind tasks by reporting current reality (i.e., a realist bias), some children fail by making phenomenist errors on reality questions (e.g., Flavell, Flavell, & Green, 1983; Frye, Zelazo, & Palfai, 1995).

Second, in light of recent evidence that test–retest reliability for false-belief tasks is rather poor (Mayes, Klin, Tercyak, Cicchetti, & Cohen, 1996), we presented children with four replications of the same task (using different materials). Like most previous studies (e.g., Freeman, Lacohée, & Coulton, 1995; Mitchell & Lacohée, 1991), Saltmarsh and Mitchell (1999) used only one test question (for

which only two answers were plausible). In addition to providing a more reliable index of children's ability, the use of multiple test questions allows assessment of the extent to which any facilitation effects are dependent on the use of particular materials (e.g., on the egg and egg box used in Freeman et al., 1995).

Finally, prior to the experiment proper, we presented children with an extensive criterial training phase designed to demonstrate that videotape provides a reliable record of past events. Saltmarsh and Mitchell (1999) also used a video training phase, but we employed a clear criterion of success, and children who did not meet this criterion were excluded from the final sample.

Thus, in Experiment 1, children were videotaped as they performed four representational change tasks, each with a different set of materials. In each task, children were shown deceptive containers, the true contents were revealed, and then children were asked standard representational change and reality questions (as a baseline measure of performance on a standard version of the task). Subsequently, children in a video support condition viewed their initial, incorrect statements on videotape, whereas children in a control condition viewed an irrelevant videotape. Finally, all children were again asked the standard representational change and reality questions. For two of the tasks, there was a 10-min delay between seeing the actual contents of the container and watching the videotape. Imposing a delay was expected to decrease the salience of the current contents of the container. According to the masking hypothesis (e.g., Mitchell, 1994), this should increase the likelihood that children will report their initial belief in response to the representational change question.

In Experiment 2, children were only asked the representational change and reality questions after the video exposure. Moreover, in this experiment, children were also asked to report their initial statement about the container (i.e., "What did you say was in the box before I opened it?").

A clear prediction that arises from work on the efficacy of mnemonic cues (e.g., Mitchell & Lacohée, 1991, and especially, Saltmarsh & Mitchell, 1999) is that children in the video support condition should report their initial incorrect beliefs because they are provided with a maximally informative reminder of those beliefs. Indeed, the videotape of their initial statement could be treated as physical evidence of their false belief, which might be expected to make children's initial belief more salient than current reality (see Mitchell, 1994). Further, if children are motivated to maintain consistency (e.g., Mitchell & Lacohée, 1991; Moses & Flavell, 1990), then the presence of an undeniable record of their initial statement might compel children to answer the representation change question correctly.

However, in light of the possibility that children could simply repeat the utterance from the videotape without understanding the test question (i.e., false positives), we reasoned that *failures* to answer the test question correctly would potentially be of greater theoretical interest than successes (cf. Freeman et al., 1995, p. 105). If 3-year-olds performed poorly on the task despite the strong mnemonic support (especially on the delay trials), this would constrain our

theoretical speculations about process models of theory of mind by suggesting boundary conditions on the facilitative effect of mnemonic cues. In addition, it would raise an important question to be addressed in future research, namely why do children fail despite such strong support? As will become clear, the current experiments did indeed find that showing children a videotape of their initial statement failed to improve their performance on a standard representational change question. Experiment 2 also demonstrated comparable results for a question asking children about their initial statements.

EXPERIMENT 1

Method

Participants

Fifty children, ranging in age from 3 years, 0 months to 4 years, 6 months (M = 3 years, 9 months; 25 girls and 25 boys), were tested. The children were recruited from local day care centers and from a database containing the names of parents who had expressed an interest in participating in research. Children were of mixed socioeconomic and ethnic backgrounds, although this information was not systematically collected. Testing took place in a quiet area at the children's day care center or in a university laboratory.

Materials

A video camera and a portable color television (15 × 18 cm) were used for taping and viewing in the training and test phases. A variety of toys, including a stuffed elephant, a Barney puppet, and a Big Bird puppet, were also employed in the training phase. In the test phase, children were presented with a total of five items. These included four misleading items: A crayon box containing a penny, a milk carton containing markers, a cereal box containing paper, and a Smarties box containing string. The fifth item was a Kleenex tissue box that actually contained tissues. This confirmation item was presented in order to confirm children's belief regarding the contents and hopefully reduce the likelihood that children would eventually expect to be deceived by the misleading items. For each item, children in the video control condition watched one of five separate video clips of an adult female who stated a single word (*moon*, *elephant*, *pillow*, *winter*, and *dinner*). Each clip had a duration of approximately 3 s (which was comparable to the duration of the videotaped statements used in the video support condition; see below).

Design and Overview of Procedure

Children were randomly assigned to a video support condition or a video control condition. Each condition consisted of a training phase and a test phase. The training phase, which was identical for both conditions, was designed to instruct children about the function of videotaping. This phase consisted of three subphases, presented in a fixed order. The first subphase emphasized visual

self-recognition, the second emphasized auditory self-recognition, and the third emphasized comprehension of the temporal dimension of video. More specifically, in the third subphase, children were shown that videotaping preserves precisely the temporal order of a videotaped event.

After training, children received a test phase in which they were presented with four misleading representational change items, the order of which was counterbalanced. Children were also presented with one nondeceptive confirmation item, which was always presented after the second representational change item (i.e., it was always presented third). For each item, children's initial statement about the item was videotaped, children were asked a representational change question and a reality question (in a counterbalanced order), and then children were shown a videotape (either immediately or after a 10-min delay). After watching the videotape, children were asked what was said on the videotape in order to assess whether they had paid attention to the videotaped information and could remember it. Then they were again asked a standard representational change question and a reality question (in a counterbalanced order). Children in the video support condition watched a videotape of their own initial statements about the item, whereas children in the video control condition watched an irrelevant videotape. Delay trials were blocked and the order of the blocks was counterbalanced (first two trials or last two trials).

Procedure

Children were tested individually by two female experimenters, a primary experimenter, who interacted with children, and a secondary experimenter, who operated the video camera. After children were comfortable with the experimenters and the setting, the training phase was initiated.

Training phase. In the first training subphase, Visual Self-Recognition, children's attention was directed to the video camera and children were asked, "Have you ever seen one of these before?" after which they were told, "The camera takes pictures of what we do. It copies us, and afterwards, we can watch what we just did." The primary experimenter was then recorded waving in front of the camera, after which she told children that "the camera just copied me, and it can go back to the beginning and show us what I just did." The tape was rewound and then played for children, who were asked Question 1, "Who is that on the television right now?" This question was asked in order to ensure that children recognized the experimenter on videotape. Viewings remained soundless in this subphase in order to emphasize visual self-recognition. After responding to Question 1, children were invited to wave in front of the camera. Prior to viewing themselves, they were asked Question 2, which was a predictive question, "Who are we going to see on television now?"[1] After answering Question 2, children watched themselves on television. Then, children were asked Questions 3 and 4, which were confirmatory questions designed to ensure that children could recognize themselves on television: "Who did we just see on television?" and "What were you doing?"

At this point, children were told, "When the camera is on, it can take pictures of us, and then we can watch what we did on television." The experimenter pointed to the television and said, "Look, the camera is on right now, so when Barney stands in front of the camera it will take his picture. Then we can see Barney on television." The experimenter demonstrated this by holding up the Barney puppet and videotaping it in front of the camera. Then, the camera was turned off and children's attention was directed toward the blank television screen. Children were told, "Look, the camera is off right now, and when the camera is off, it can't take our picture and we won't get to see it on television." Children were then told, "Now Big Bird will stand in front of the camera because he wants to take his picture. But remember, the camera is off, and so Big Bird doesn't get to take his picture this time, and he won't be on television." Children were then asked a predictive question (Question 5), "Who are we going to see on television now, Barney or Big Bird?" If children responded incorrectly (i.e., if they said, "Big Bird") then the sequence was retaped and they were asked the question again. The purpose of this subphase was to demonstrate that the camera can record specific visual events and that these events can be reviewed later.

An analogous demonstration was included in the second subphase, Auditory Self-Recognition, in which children were told that the camera can "hear everything we say, and then we can listen to what we just said on television." The experimenter made one brief statement with the camera turned off, "I like cats," and one with the camera turned on, "I like dogs." Children were then asked, "Am I going to say that I like cats or am I going to say that I like dogs?" (Question 6). After watching the videotape, children were asked a confirmatory question, "What did I just say on television?" (Question 7), to ensure that they could use the television as a source of information about what had been said in the past. Subsequently, children were invited to make two statements, one (regarding their favorite color) with the camera turned on and one (regarding their favorite animal) with the camera turned off. They were then asked a predictive question, "Tell me what we are going to hear on television now, are you going to say your favorite color or are you going to say your favorite animal?" (Question 8), and, after watching the videotape, they were asked another confirmatory question, "On TV, what did you say?" (Question 9). In cases where children responded incorrectly to these questions, they received a verbal correction. If children responded incorrectly to either Question 6 or 8, then that sequence was retaped, and children were asked the question again. The purpose of this subphase was to demonstrate

[1] Based on previous findings using photographs (Zaitchik, 1990), one might expect 3-year-olds to find Question 2 (and Questions 5, 6, and 8; see below) confusing. In Zaitchik's task, children were shown a situation, a photograph was taken of the situation, the situation was changed, and children were asked to predict the content of the photograph. Three-year-olds had difficulty with this question. However, Slaughter (1998) has recently found that 3-year-olds perform well on Zaitchik's task and that children benefit from training. Moreover, in the current experiment, the questions of interest during the test phase (Video-Say and representational change; see below) are arguably more analogous to the confirmatory questions than to the predictive questions.

(for self and for other) that the camera can record specific statements that can be listened to later.

Finally, in the third subphase, Temporal Understanding, children were introduced to a battery-operated toy, Elmo the elephant, and told, "We're going to watch Elmo walk." The experimenter emphasized that the camera would be watching Elmo and copying everything he did, from where he started walking to where he stopped walking. Children watched the toy walk for approximately 10 s, during which their attention was directed to the live recording of the elephant moving from the left to the right side of the television screen. At this point, taping was conspicuously stopped, and the videotape was paused in front of children. The experimenter told the children, "We can watch what Elmo did all over again," and informed them, "The camera can do something special; it can show Elmo walk backwards, back to where he started. Let's watch him go back to the beginning."

The experimenter rewound the tape in front of children and pointed out, "We can watch Elmo go back to the beginning." When the toy reached the left side of the screen, the tape was paused, children were told that the tape would be played, and then they were asked Question 10, which was predictive: "Which way is Elmo going to go now (to go forward)?" Then, the tape was played and children watched the toy move to the other side of the television screen. Afterward, children were asked Question 11: "Which way does Elmo go to get back to the beginning (to go backwards)?" Children were directed to the television screen to demonstrate their answer. Again, children who responded incorrectly to a question were corrected. The purpose of this subphase was to demonstrate, in a very transparent fashion, that the video camera preserves the temporal order of events.

Children received a score ranging from 0 to 11 points for the training phase, with 1 point awarded for each question that children answered correctly on the first attempt. In order to pass, children were required to reach a criterion of eight questions (73%) correct, or four of the last five questions correct. The probability of answering eight or more questions correctly based on random responding (and assuming that there were two plausible answers to each question) is $p < .12$ (binomial theorem), whereas the probability of answering four or five of five is $p < .19$.

Test phase. In the video support condition, children were told, "We're going to play some games now. The camera is on; remember that it copies what we do and say, and later we can listen to it." Then children received five trials, including four misleading representational change items and one confirmation item. For the misleading items, children were shown a familiar container (e.g., a Smarties box) and asked: "What's this?" Children were videotaped during this interaction. Then, after children's initial statement about the item, videotaping was conspicuously stopped, the unexpected contents of the container (e.g., string) were revealed, and children were asked a prevideo representational change question, "What did you think was in the box before I opened it up, ___ or ___?" and a prevideo reality question, "What's really in the box,

___ or ___?" (the correct–incorrect order of the forced-choice options was alternated).

For two of the misleading items, children were then presented immediately with their initial videotaped statement (e.g., saying "Smarties") about the container. The videotaped segment lasted about 3 s and consisted of the experimenter asking children, "What is this?" in the presence of the unopened container, followed by children's response. For the remaining two misleading items, there was a 10-min delay (during which children participated in an unrelated task) before children were exposed to the videotape. In all cases, the experimenter rewound the tape and told children, "We're going to listen to what we said about this box when we first saw it." After viewing the videotape, all children were asked a Video-Say question, "On television, what did you say?" followed by postvideo representational change and reality questions. (The Video-Say question was asked in an open-ended fashion in order to make it comparable to that used in the video control condition, for which no plausible pair of options was available).

The procedure for the confirmation item (i.e., the nondeceptive Kleenex box) was identical to that for the misleading items, except that no forced choice was given for any questions. Performance on this item was not analyzed because nearly all children in both conditions (36 of 38 children) answered both the "representational change" question[2] and the reality questions correctly for this item. The two children who failed to do so (one child in each condition) either refused to respond to one or more of the questions or stated that they did not know the answer.

Children in the video control condition received essentially the same procedure as children in the video support condition. However, instead of viewing their own videotaped behavior, children in the video control condition watched a series of five irrelevant video clips in each of which a woman stated a single word (*moon*, *elephant*, *pillow*, *winter*, or *dinner*). Instructions referring to the videotape were changed accordingly. Also, because they did not view their own statements statements on videotape, they were asked a Video-Say question that referred to the woman on the videotape, "On television, what did Laura say?"

Results

A total of 50 children were tested. However, 7 children (including 3 boys in the video support condition, 1 boy in the video control condition, and 3 girls in the video control condition) refused to cooperate during the training phase. Data from 1 additional boy (in the video support condition) were not included due to problems with the testing equipment. Finally, 4 boys (2 in the video support condition and 2 in the control condition) failed the training phase (see below) and were excluded from further data analyses. Of the remaining 38 children, there were 21 in the video support condition (8 boys and 13 girls) and 17 in the video control condition (8 boys and 9 girls). A one-way analysis of variance (ANOVA)

[2] Note that children's representations regarding the confirmation item did not actually change.

confirmed that the mean age of children in the video support condition ($M = 45.9$ months, $SD = 5.06$) did not differ significantly from that of children in the video control condition ($M = 44.8$ months, $SD = 5.50$), $F(1, 36) < 1$, *ns*.

Performance on the Training Phase

As expected, performance on the training phase did not differ significantly across conditions, $F(1, 36) < 1$, *ns*. Children in the video support condition answered a mean of 8.89 ($SD = 1.17$) training questions correctly (i.e., 81%) and children in the video control condition answered a mean of 8.62 ($SD = 1.2$) questions correctly (78%). Most of the children who reached criterion (33 of 38) did so by answering 8 or more of 11 questions correctly ($p < .12$). Analysis of performance on individual training questions revealed that performance varied across questions, $F(10, 370) = 24.64$, $p < .0001$, and that Questions 5 ($M = .39$, $SD = .5$), 6 ($M = .45$, $SD = .5$), and 8 ($M = .39$, $SD = .5$) were especially difficult. Mean scores on all other training questions were greater than .81.

Baseline Performance on the Prevideo Representational Change and Reality Questions

All children received prevideo representational change and reality questions for each of four items. Children received a score of 1 for each item for which they answered both the representational change and reality questions correctly; otherwise, children received a score of zero for that item. An initial inspection of the data indicated that these scores (referred to as representational change plus reality scores) were distributed across the full range from 0 to 4. In order to confirm that performance in the two conditions did not differ prior to any intervention, a one-way ANOVA was conducted on prevideo test performance. This ANOVA indicated that baseline performance in the video support condition did not differ significantly from baseline performance in the video control condition, $F(1, 36) < 1$, *ns* (see Table 1 for the means).

In the majority of cases where children erred (72 of 119 errors; 60.5%), children did so by answering the representational change question in terms of current reality while answering the reality question correctly (i.e., they made realist errors). There were 22 cases (18.49%) where children answered the representational change question correctly but answered the reality question on the basis of their initial belief (i.e., phenomenist errors) and 24 cases where children's responses to the two questions were reversed. (A remaining child answered the representational change question but said, "Nothing," in response to the reality question.)

Performance on the Postvideo Representational Change and Reality Questions

Performance on the postvideo representational change and reality questions was scored in the same fashion as performance on the corresponding prevideo questions. Again, inspection of the data indicated that representational change plus reality scores were distributed across the full range from 0 to 4. To assess the

TABLE 1
Mean Number (and Standard Deviation) of Pre- and Postvideo Test Questions Answered Correctly in Experiment 1

Condition	Question		
	Prevideo RC+RE	Video-Say	Postvideo RC + RE
Video support (n = 21)	.95 (1.50)	2.95 (.97)	.90 (1.58)
Video control (n = 17)	.76 (1.25)	3.71 (.45)	1.0 (1.54)

Note. Means are of a possible score of 4. RC + RE = representational change plus reality scores.

effects of video support and delay on children's ability to recall their previous beliefs, these scores were submitted to a Condition (video support vs video control) × Delay (immediate vs delay) mixed ANOVA, with Delay as a within-subject variable. This ANOVA produced no significant effects or interactions. Indeed, none of the effects or interactions even approached significance. In particular, there were no main effects of condition, $F(1, 36) < 1$, *ns* or delay, $F(1, 36) < 1$, *ns*. The effect sizes associated with these comparisons were very small (Cohen's *d*s were .02 and .16, respectively; Cohen, 1988). Table 1 presents the means for each condition, collapsed across delay because condition was the variable of primary interest.

As on the baseline prevideo measures, in the majority of cases where children erred (78 out of 121 errors; 64.46%), children made realist errors. The remaining cases consisted of the following: 20 phenomenist errors (16.53%), 22 cases in which responses to the two questions were reversed (18.18%), and 1 case in which a child said, "Nothing," in response to the reality question.

Performance on the Video-Say Question

The Video-Say question was designed to determine whether children paid attention to the information that was presented to them on videotape and were able to remember this information. Children were given 1 point for each Video-Say question that they answered correctly. Because there was no corresponding reality question, performance on this question was compared to performance on the representational change question alone (as opposed to representational change plus reality scores). A Condition x Question (Video-Say vs representational change) mixed ANOVA with question as the within-subject variable showed only that children performed significantly better on the Video-Say question (M = 3.29 of 4, SD = .87) than they did on the representational change question (M = 1.50 of 4, SD = 1.66), $F(1, 36) = 44.37$, $p < .0001$. See Table 1 for means according to condition. (A similar analysis comparing scores on the Video-Say question to representational change plus reality scores revealed an identical pattern of results.) When children erred on the Video-Say question, they almost invariably did so by giving realist answers. In the video support condition, where the large majority of errors occurred (22 of 27), 19 (86.36%) of the errors were realist errors (two children said they did not know and one child answered, "Nothing").

Correlation Analyses

Pearson correlation coefficients were calculated for training scores, pre- and postvideo representational change plus reality scores, and Video-Say scores. The only significant correlations to emerge were between training scores and Video-Say scores, $r(38) = .34$, $p < .05$, and between prevideo representational change plus reality scores and postvideo representational change plus reality scores, which were very highly correlated, $r(38) = .84$, $p < .0001$.

Categorical Analyses of Performance

Children's performance during the test phase was also scored categorically. Based on the results of the quantitative analyses, and in order to support comparisons against chance responding, data were collapsed across levels of delay. In order to pass each of the individual questions (i.e., pre- and postvideo representational change and reality questions, considered singly, and the Video-Say question), children were required to respond correctly on three or four of the four items ($p = .31$, based on random responding). The same criterion was used for pre- and postvideo representational change plus reality questions, in which case the probability of reaching the criterion by chance was $p < .05$. The numbers of children passing and failing each question (and percentages) according to this criterion are presented in Table 2. As can be seen, more children in each condition passed the Video-Say question than would be expected by chance (based on a second-order application of the binomial distribution, $p < .05$), whereas more children in each condition failed all of the other questions, $p < .05$). The pattern of results was extremely clear, obviating the need for additional formal analyses: Children performed better on the Video-Say question than they did on the postvideo representational change plus reality questions, the video intervention had no impact on children's performance (pre- vs postvideo scores), and there was no effect of condition.

Discussion

Experiment 1 was designed to assess the effect of video reminders on 3-year-olds' recollection of their previous beliefs in a representational change paradigm. All children who were included in the final sample performed well on a criterial

TABLE 2
Numbers of Children (and Percentages) in Each Condition Passing Pre- and Postvideo Test Questions in Experiment 1

	Question		
Condition	Prevideo RC+RE	Video-Say	Postvideo RC + RE
Video support ($n = 21$)	5 (24%)	15 (71%)	4 (19%)
Video control ($n = 17$)	2 (12%)	17 (100%)	3 (18%)

Note. RC + RE = representational change plus reality scores.

training task designed to demonstrate that a video camera can be used to create a reliable record of their behavior. In the experiment proper, children in the video support condition were shown a videotape of their initial incorrect statement about a misleading container (either immediately or after delay). These children were successful at reporting their initial statement in response to the Video-Say question, "On television, what did you say?," which provides evidence that their behavior was influenced by the information presented to them on the videotape; children did not fail to attend to the cue, nor did they forget the information that it contained. Remarkably, however, these children performed just as poorly on the standard representational change question as children in the control condition (who watched an irrelevant videotape), and their performance was no better *after* seeing the videotape than it was before seeing it. Indeed, in both the video support and the video condition, the large majority of children answered both the pre- and postvideo representational change questions incorrectly, and pre- and postvideo performance was highly correlated. Across the four different items, children typically provided a realist answer for all but the Video-Say question, to which they responded with their initial statement. This pattern of responding, together with the high correlation between pre- and postvideo scores, renders it unlikely that children's errors can be attributed to their misunderstanding of the pragmatic implications of repeated questions (pace Siegal, 1997)—for example, children did not change their responses on successive questions.

Children's poor performance in the video support condition was perhaps especially striking on the delay trials. In light of the masking hypothesis (see Mitchell, 1994), it was expected that imposing a 10-min delay between seeing the actual, unexpected contents of the container and seeing the videotape of the initial statement might reduce children's bias to respond on the basis of reality (perhaps by reducing the salience of children's knowledge of the actual contents). However, in this experiment, performance did not differ significantly as a function of delay—children performed poorly despite the delay.

The findings from the current study are in contrast to earlier findings (e.g., Mitchell & Lacohée, 1991) that indicated that children can report their initial, incorrect beliefs in the representational change paradigm when they are provided with salient (physical) reminders of those beliefs. In particular, the results are in contrast to the findings from Saltmarsh and Mitchell (1999), who found that video reminders improved 3-year-olds' performance on representational change questions. If the findings from Experiment 1 are reliable, then they point to limits on the facilitation of false belief recollection through the provision of mnemonic cues. Consequently, the findings would constrain process modeling of performance on representational change tasks, and they may prompt a reinterpretation of previous demonstrations of facilitation.

In Experiment 2, we attempted to assess the reliability of our findings, but we also sought to explore an additional issue. Children may have been obliged by the videotaped "evidence" to acknowledge what they had said, even though they refused to admit (or perhaps believe) that they had actually believed it. If so, it

follows that children would succeed on a Referential-Say question that referred to the misleading object ("What did you say was in the box before I opened it?"), but not to the videotape. On the other hand, if children perform well on the Video-Say question but poorly on a Referential-Say question, then perhaps on the Video-Say question they simply repeat the initial statement from the videotape without any understanding of how it relates to the earlier episode.

There were five modifications to the procedure in Experiment 2 in addition to the creation of a second Say question. First, in order to compare children's performance on the test questions, we attempted to make the syntactic structure of the different questions as similar as possible. Thus, the Video-Say question was changed to, "What did you say was in the box on television?" in order to make it more comparable to the representational change ("What did you think was in the box before I opened it?") and Referential-Say questions ("What did you say was in the box before I opened it?"). Second, because the Video-Say question was originally asked in an open-ended form, all test questions (representational change, reality, Video-Say, and Referential-Say) were asked in this manner in Experiment 2. Third, because Experiment 1 revealed that there was no difference in performance on the first and second set of representational change and reality questions, children were asked these questions only once, after the video exposure. Fourth, because there was no effect of delay in Experiment 1, all children were exposed to the videotape immediately after seeing the actual contents of the container. Fifth, there was no control condition in this experiment, as we were interested in comparing children's responses on the different types of question.

EXPERIMENT 2

Method

Participants. A total of 27 children were recruited for Experiment 2. However, 6 children (3 girls and 3 boys) refused to cooperate with the experimenter. In addition, 1 girl failed to reach criterion in the training phase (scored as in Experiment 1) and complete data from another girl could not be gathered due to problems with the video equipment. The final sample consisted of 19 children ranging in age from 3 years, 3 months to 4 years, 0 months (M = 3 years, 7 months; 10 boys, 9 girls). Participants were recruited in the same manner as Experiment 1, although no child participated in both experiments.

Design and procedure. All children received a training phase (identical to that employed in Experiment 1) and a test phase. In the test phase, children's initial statements about four misleading representational change items (presented in a counterbalanced order) and one confirmation item (always presented third) were videotaped. The items were the same as those used in Experiment 1. For each item, after children stated their initial belief about the contents of the container, the actual contents were revealed. Then, children proceeded immediately to watch the videotape of their initial statement about the item. At this point, children were asked a Video-Say question, which was followed by a representational change

question, a reality question, and a Referential-Say question, presented in a counterbalanced order. The Video-Say question was, "What did you say was in the box on television?"[3] The representational change and reality questions were identical to those used in Experiment 1 with the exception that they were presented as open-ended questions. The Referential-Say question was, "What did you say was in the box before I opened it?" Table 3 lists each test question together with a brief description of its purpose in the experiment.

Results

Performance on the training phase. Performance on the training phase was scored in a manner identical to that of Experiment 1. The mean training score for participants in the final sample was 8.16 (SD = 1.26). Most of the children who reached criterion (15 of 19) did so by answering 8 or more of 11 questions correctly ($p < .12$). Analysis of performance on individual training questions revealed that performance varied across questions, $F(10, 180) = 11.06$, $p < .0001$, and that the predictive Questions 2 ($M = .58$, $SD = .51$), 5 ($M = .32$, $SD = .48$), 6 ($M = .53$, $SD = .5$), and 8 ($M = .26$, $SD = .45$) were especially difficult. Mean scores on all other training questions were greater than .79.

Performance on the postvideo test questions. All children answered all questions concerning the confirmation item correctly, except for one boy who answered, "I don't know," in response to the Referential-Say question for that item. The effect of question on children's performance on the misleading items was initially assessed by comparing performance on the Video-Say, Referential-Say, representational change, and reality questions, with the last two questions

TABLE 3
Summary of Questions Presented in Experiment 2

Question type	Question	Purpose
Video-Say	What did you say was in the box on television?	Assessed memory for the cue
Referential-Say	What did you say was in the box before I opened it?	Assessed memory of initial statement without provision of the cue
Representational change	What did you think was in the box before I opened it?	Assessed memory of initial belief without provision of the cue (standard question)
Reality	What's really in the box?	Assessed memory of true content of container

[3] One reviewer pointed out that the Video-Say question in Experiment 2 was ambiguous. That is, children might have interpreted the question to be asking what they said about the box that they saw on television rather than what they said on television about the box. The former interpretation is potentially closer to the Referential-Say question than it is to the unambiguous Video-Say question used in Experiment 1. As will become clear, however, children typically answered the question correctly, as they did the Video-Say question in Experiment 1. In contrast, they had difficulty with the Referential-Say question. These findings suggest that children interpreted the question as intended.

considered singly so that the probability of success based on random responding would be equated across all four questions. (The reality question was also included singly in this analysis in order to determine whether performance on each SAY question was more like performance on the representational change question or on the reality question). Inspection of the data indicated that scores were distributed across the entire range, even for the reality question. A one-way ANOVA revealed a main effect of question, $F(3, 54) = 12.54$, $p < .0001$. Tukey's HSD tests determined that performance on the Video-Say ($M = 3.00$, $SD = 1.05$) and reality questions ($M = 2.95$, $SD = 1.31$), which did not differ, was better than performance on the Referential-Say ($M = 1.42$, $SD = 1.46$) and representational change questions ($M = 1.37$, $SD = 1.34$), which did not differ. The same pattern of results was found when Say scores were compared to representational change plus reality scores ($M = 1.0$, $SD = 1.33$), which were generated as in Experiment 1 (i.e., Referential-Say resembled representational change plus reality).

There were 48 cases where children answered the representational change question incorrectly, and in all cases, children did so by giving realist answers. Children also made realist errors on the Video-Say question (18 of 19 errors; 94.74%) and the Referential-Say question (44 of 48 errors; 91.67%). When children erred on the reality question (20 cases), they always did so by repeating their initial statement about the contents of the container (phenomenist answers).

Correlational analyses. Pearson correlation coefficients were calculated for training scores and postvideo test scores. The only significant correlations to emerge were between Referential-Say and representational change, $r(19) = .82$, $p < .0001$, and between Referential-Say and representational change plus reality scores, $r(19) = .68$, $p < .01$.

Categorical analyses of performance. Children's performance during the test phase was also scored categorically as in Experiment 1 (i.e., in order to pass, children were required to respond correctly on three or four of the four items). According to this criterion, 13 children (68%) passed the Video-Say question and 14 (74%) passed the reality question. In contrast, only 3 children (16%) passed the representational change question and only 4 children (21%) passed the Referential-Say question. Two children (10%) passed the representational change plus reality questions considered together. More children passed the Video-Say and reality questions than would be expected by chance (based on the binomial distribution, $p < .05$), whereas more children failed all of the other questions, $p < .05$.

Discussion

The results of Experiment 2 replicated the finding from Experiment 1 that children performed well on the Video-Say question and yet failed to answer the representational change question correctly. Moreover, in Experiment 2, children also performed poorly on a Referential-Say question that referred to the misleading object, but not to the cue, "What did you say was in the box before I opened it?" Children's performance on the Referential-Say question was

significantly worse than their performance on the Video-Say question; indeed, children performed as poorly on the Referential-Say question as they did on the representational change question.

Two main conclusions can be drawn from this experiment. First, these results are inconsistent with the notion that children felt obliged by the videotaped "evidence" to acknowledge what they said despite denying what they believed; if that were the case, then they should have responded correctly on the Referential-Say question. Second, children's poor performance on the Referential-Say question indicates that on this task, children do not find it easier to report previous statements than previous beliefs.

GENERAL DISCUSSION

In two experiments, children were shown a videotape of their initial incorrect statement about a misleading container prior to being asked standard representational change and reality questions. In both experiments, children in the video support condition were able to report their initial statement about the container when asked what they said on television. This finding is important because it ensures that children paid attention to, and remembered, the information that was provided on the videotape. Remarkably, however, in Experiment 1, the overwhelming majority of children nonetheless failed to report their initial belief about the container. Imposing a time delay between revelation of the contents and exposure to the videotape (in an effort to reduce the salience of children's knowledge about the true contents) had little effect of children's performance. Experiment 2 replicated the main findings from Experiment 1 and also found that children failed to report their initial statement about a misleading item when the question referred specifically to that item but did not refer to the videotape. This latter finding indicates that children were not forced by the strong videotaped evidence to acknowledge what they had said (perhaps while refusing to admit that they had held a false belief). The finding also indicates that it was not simply easier in this task for children to answer questions about statements than beliefs. Together, the findings from Experiments 1 and 2 constrain our interpretation of the facilitative effect of mnemonic cues on recollection of false beliefs and pose an interpretive challenge to be addressed in future research.

Although the current findings are surprising, they are broadly consistent with previous findings by Robinson, Riggs, and their colleagues. First, in a study by Robinson et al. (1996), children drew a picture that corresponded to their initial belief about the contents of a misleading container. The majority of children were able to recall what they had drawn, but they failed to use this information to answer a representational change question correctly. Second, Riggs and Robinson (1995, Experiment 3) investigated whether children could use statements as cues to help them infer another character's false belief. Two conditions were compared in which a character left a room while a red mug was switched with a white one. In one condition, when the character returned, he stated that he thought the mug

was red. In another condition, there was no verbal cue. Children performed equally poorly on a false belief question in both conditions; there was no evidence that the statement cued them to infer the character's false belief. As in the study by Robinson et al. (1996), then, a strong, explicit cue failed to improve children's performance on a task assessing reasoning about beliefs.

Although the findings from the current experiments are consistent with those of Robinson and Riggs (e.g., Robinson et al., 1996), the findings contrast with demonstrations that children benefit from reminders on the representational change task (e.g., Mitchell & Lacohée, 1991; Saltmarsh & Mitchell, 1999). In particular, our results contrast with the results of Saltmarsh and Mitchell (1999). In that study, children were more likely to respond correctly on several types of questions when they viewed a videotape of their initial statements than when they did not. In two experiments, the authors used cued questions that referred to the camera when asking children to recall their initial statements, and no child received these questions together with reality questions. However, in Experiment 2, children in one group were asked a standard representational change question ("When I first showed you the box, before we opened it up, what did you think was inside?") that was probably comparable to the ones employed in the current experiments ("What did you think was in the box before I opened it up?"). Saltmarsh and Mitchell found that video cues improved performance on this question, too.

In the absence of additional research, it is impossible to determine why our results differed from those of Saltmarsh and Mitchell (1999). Of course, the discrepancy may be due to the sort of unreliability of empirical findings that has been noted by other authors (e.g., by Mitchell, who was twice unable to replicate the results of Mitchell & Lacohée, 1991; see Freeman & Lacohée, 1995, p. 55; or by Robinson & Mitchell, 1994, who were unable to replicate one of the findings from Robinson & Mitchell, 1992). Even if it is not, however, we may need to reinterpret previous demonstrations of the efficacy of mnemonic cues (and contextual support in general). For example, it is possible that observed facilitation of performance in many previous studies was due to the use of nonstandard test questions. Mitchell and Lacohée (1988) asked children, "When you posted your picture in the postbox (indicating the closed postbox), what did you think was in here (indicating the closed test item)?" Perhaps children tended to answer this question correctly because they were referring to the content of the picture, without fully appreciating its relevance to what they previously believed. Mitchell and Lacohée argued against this interpretation because children in an irrelevant posting condition did not answer the modified representational change question on the basis of the irrelevant picture in the postbox. Freeman and Lacohée (1995) also argued against this interpretation because they found that different cues (e.g., a picture of an egg vs an actual hardboiled egg) were differentially effective even though they shared the same name. The possibility remains, however, that the different cues simply varied in the likelihood that they would encourage children to name them (as cues) rather than in their efficacy as mnemonic cues. In the Mitchell and Lacohée experiments, for

example, it is possible that the irrelevant alternatives (pictures of cartoon characters) were ignored when the test question was asked precisely because they were implausible options to the question, ". . .What did you think was in here?" On the other hand, the relevant cues were both (a) present at posting and (b) plausible options to the test question. To account for Freeman and Lacohée's (1995) finding that pictures of the content of a false belief (e.g., a picture of an egg) are more effective cues than samples (e.g., an actual hardboiled egg), it would be necessary to claim that children are more likely to label pictures than samples. Given the use to which pictures (vs familiar objects) are frequently put in childhood routines (e.g., Whitehurst, Arnold, Epstein, Angell, Smith, & Fischel, 1994) this claim seems reasonable. Unfortunately, however, this explanation does not address the discrepancy between the current findings and the finding from Saltmarsh and Mitchell (1999, Experiment 2), which used a noncued question. As noted, further research will be needed to determine whether Saltmarsh and Mitchell's finding is reliable and, if so, how to explain the discrepancy.

The hypothesis that children answer cued representational change questions correctly because children refer to the cue rather than because the cue reminds them of the previous belief fits neatly with Robinson et al.'s (1996) finding that children who drew the supposed contents of a container could correctly recall the content of the drawings, but still failed to report their initial incorrect belief. The hypothesis also accounts for the finding from the current experiment that children performed well on the Video-Say questions despite performing poorly on the representational change questions.

If children did understand the implications of the Video-Say question for the representational change question, then they failed to act on the basis of this understanding and they did not appear to be bothered by the inconsistency in their responses. Such dissociations between knowing and doing have been observed in children in this age range. For example, in a study by Zelazo, Frye, and Rapus (1996), 3-year-old children who were asked to sort cards first by one dimension (e.g., color) and then by another (e.g., shape) blithely persisted in sorting by the first dimension despite being told the new rules on every trial and despite answering questions about the new rules. In that study, as in the current experiments, children's responses to particular test questions were rather narrowly dependent on the local context of the questions, and children apparently failed to detect inconsistencies between their responses to different questions (cf. Inhelder & Piaget, 1964; Vygotsky, 1978).

Several considerations support the suggestion that children did understand that videotape provides a reliable record of the past, even though they failed to respond correctly to the representational change questions. First, children received an extensive training session designed to demonstrate visual and auditory recording as well as the temporal aspects of taping. In the experiment proper, children needed to recognize themselves and the experimenter on television (corresponding to training questions 1, 3, and 4 in the visual self-recognition subphase), repeat auditory information given to them on television

(questions 7 and 9 in the auditory self-recognition phase), and understand that the camera records an event in a way the preserves its temporal order (Temporal Understanding subphase). The great majority of children performed very well on these task-relevant questions. Thus, any difficulty that children had with the predictive questions on the training phase (questions 2, 5, 6, and 8) seems unable to account for the pattern of results obtained from the test phase. Another consideration is that even in the absence of an adequate understanding of video, children might have been able to make the appropriate inference on the basis of the verbal information provided by the experimenter (together with the videotaped information). That is, on each trial, prior to watching the television, children in the video support condition were told, "We're going to listen to what we said about this [item] when we first saw it." To answer the representational change question correctly, children needed simply to repeat what was said on the videotape, even if they did not fully understand its relation to the past event.

Nonetheless, it is possible children's difficulty on the representational change question can be attributed in part to a fragile understanding of the implications of video—perhaps in combination with other limitations. Research on children's understanding of external representations more generally has revealed that this understanding develops gradually during the preschool years (Liben, 1999), and there is no reason to believe that the development of video understanding is peculiar. Thus, although 2.5-year-olds are able to use video information in some circumstances (Troseth & DeLoache, 1998), even 3-year-olds have difficulty using it in other circumstances. Indeed, consistent with work using other types of external representation (e.g., photographs; Zatichik, 1990; and drawings; Thomas, Jolley, Robinson, & Champion, 1999), Zelazo, Sommerville, and Nichols (1999) found that 3-year-olds have difficulty appreciating the implications of delayed videotaped information when that information is in conflict with children's initial belief. More specifically, these authors used a variant of Troseth and DeLoache's (1998) search task and found that 3-year-olds had difficulty using videotaped information to guide their search for an object hidden when they had a prior expectation about the object's location. These authors argued that 3-year-olds have particular difficulty using delayed (vs live) external representations (including video) to guide their behavior when the information provided by these representations conflicts with a salient alternative (e.g., reality). When provided with conflicting information via a delayed and dimly understood representational medium such as video, the conflicting information may be ignored or treated as somehow irrelevant to the situation. A difficulty of this type could well account for the failure to find facilitation in the current experiments.

Difficulty using information that conflicts with initial beliefs could also account for the finding that in both of the current experiments, children in the video support conditions only answered an average of about three of four Video-Say questions correctly. In contrast, children in the video control conditions performed at ceiling. Again, it is possible that children did not fully

comprehend this question, but it seems likely *in any case* that the true contents of the container were so salient to children that sometimes they simply disregarded their initial statement on television, even when responding to the Video-Say question.

Presumably, if children's poor performance on standard representational change tasks were due to a memory limitation, and not a more profound deficit, then providing them with a retrieval cue such as a videotape of their initial statement would have facilitated recollection of the initial belief or at least helped them to reconstruct that belief regardless of whether they completely understood the medium through which the cue was presented. The findings from Experiments 1 and 2 failed to reveal facilitation, despite demonstrating that children were able to report the statement that appeared on video. Together with other findings (e.g., Robinson et al., 1996), these findings place boundary conditions on previous demonstrations of facilitation and pose an interesting challenge for future research. If 3-year-olds can remember and report what they said on the videotape, why do they fail to report their initial statement? One possibility is that 3-year-olds have difficulty using *videotaped* information, in particular, to inform their responding when the information provided by videotape conflicts with their current belief (Zelazo et al., 1999). Another possibility is that 3-year-olds' difficulty with the representational change task is so severe that they require considerable environmental support to overcome it (i.e., more support than was provided by the video reminder in the current experiments). In either case, several accounts of children's difficulty with this task (and other theory of mind tasks) remain viable options, including (a) the possibility that children fail to retain information about previous beliefs (Gopnik & Graf, 1988); (b) the possibility that 3-year-olds fail to comprehend the concept of belief (e.g., Astington & Gopnik, 1988; Perner, 1991), rendering the test questions senseless; (c) the possibility that children have difficulty reasoning counterfactually (Riggs, Peterson, Robinson, & Mitchell, 1998); and (d) the possibility that children have difficulty integrating multiple representations of a single thing (e.g., DeLoache & Burns, 1993; Flavell, Flavell, & Green, 1983; Frye et al., 1995).

REFERENCES

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

Astington, J. W., & Gopnik, A. (1988). Knowing you've changed your mind: Children's understanding of representational change. In J. W. Astington, P. L. Harris, & D. R. Olson (Eds.), *Developing theories of mind* (pp. 193–207). Cambridge, MA: Cambridge Univ. Press.

Bjorklund, D. F. (2000). *Children's thinking: Developmental function and individual differences* (3rd ed.). Belmont, CA: Wadsworth.

Carlson, S. M. (1997). *Individual differences in inhibitory control and children's theory of mind*. Paper presented at the biennial meeting of the Society for Research in Child Development, April, 1997, Washington, DC.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Davis H. L., & Pratt, C. (1995). The development of children's theory of mind: The working memory explanation. *Australian Journal of Psychology,* **47**, 25–31.

DeLoache, J. S., & Burns, N. M. (1993). Symbolic development in young children: Understanding models and pictures. In C. Pratt & A. F. Garton (Eds.), *Systems of representation in children: Development and use* (pp. 91–112). London: Wiley.

Flavell, J. H., Flavell, E. R., & Green, F. L. (1983). Development of the appearance–reality distinction. *Cognitive Psychology,* **15**, 95–120.

Freeman, N. H. (1994). Associations and dissociations in theories of mind. In C. Lewis & P. Mitchell (Eds.), *Children's early understanding of mind* (pp. 95–111). Hove, UK: Erlbaum.

Freeman, N. H., & Lacohée, H. (1995). Making explicit 3-year-olds' implicit competence with their own false beliefs. *Cognition,* **56**, 31–60.

Freeman, N. H., & Lacohée, H., & Coulton, S. (1995). Cued-recall approach to 3-year-olds' memory for an honest mistake. *Journal of Experimental Child Psychology,* **60**, 102–115.

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

Gopnik, A. (1993). How we know our minds: The illusion of first-person knowledge of intentionality. *Behavioral and Brain Sciences,* **16**(1–15), 90–101.

Gopnik, A., & Astington, J. W. (1988). Children's understanding of representational change and its relation to the understanding of false belief and the appearance–reality distinction. *Child Development,* **59**, 26–37.

Gopnik, A., & Graf, P. (1988). Knowing how you know: Young children's ability to identify and remember the sources of their beliefs. *Child Development,* **59**, 1366–1371.

Gordon A. & Olson, D. (1998). The relation between acquisition of a theory of mind and information processing capacity. *Journal of Experimental Child Psychology,* **68**, 70–83.

Inhelder, B., & Piaget, J. (1964). *The early growth of logic in the child* (E. A. Lunzer & D. Papert, Trans.). New York: Harper & Row.

Keenan, T., Olson, D. R., & Marini, Z. (1998). Working memory and children's developing theories of mind. *Australian Journal of Psychology,* **50**, 76–82.

Liben, L. S. (1999). Developing an understanding of external spatial representations. In I. E. Sigel (Ed.), *The development of mental representations: Theories and applications* (pp. 297–321). Mahwah, NJ: Erlbaum.

Mayes, L. C., Klin, A., Tercyak, K. P., Cicchetti, D. V., & Cohen, D. J. (1996). Test–retest reliability of false-belief tasks. *Journal of Child Psychology and Psychiatry,* **37**, 313–319.

Mitchell, P. (1994). Realism and early conception of mind: A synthesis of phylogenetic and ontogenetic issues. In C. Lewis & P. Mitchell (Eds.), *Children's early understanding of mind* (pp. 19–45). Hove, UK: Erlbaum.

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

Moses, L. J., & Flavell, J. H. (1990). Inferring false beliefs from actions and reactions. *Child Development,* **61**, 929–945.

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

Perner, J., & Lang, B. (1999). Development of theory mind and executive function. *Trends in Cognitive Sciences,* **3**, 337–344.

Perner, J., Stummer, S., & Lang, B. (1999). Executive function and theory of mind: Cognitive complexity or functional dependence? In P. D. Zelazo, J. W. Astington, & D. R. Olson (Eds.), *Developing theories of intention: Social understanding and self-control* (pp. 133–152). Mahwah, NJ: Erlbaum.

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

Riggs, K. J., & Robinson, E. J. (1995). What people say and what they think: Children's judgement of false belief in relation to their recall of false messages. *British Journal of Developmental Psychology,* **13**, 271–284.

Robinson, E. J., Riggs, K. J., & Samuel, J. (1996). Children's memory for drawings based on a false belief. *Developmental Psychology,* **32**, 1056–1064.

Saltmarsh, R., & Mitchell, P. (1999). Indelible evidence of false belief: Confronting young children with video recordings of themselves. *Psychologia,* **42**, 145–159.

Siegal, M. (1997). *Knowing children: Experiments in conversation and cognition* (2nd ed.). Hove, UK: Psychology Press.

Slaughter, V. (1998). Children's understanding of pictorial and mental representations. *Child Development,* **69**, 321–332.

Thomas, G. V., Jolley, R. P., Robinson, E. J., & Champion, H. (1999). Realist errors in children's responses to pictures and words as representations. *Journal of Experimental Child Psychology,* **74**, 1–20.

Troseth, G. L., & DeLoache, J. S. (1998). The medium can obscure the message: Young children's understanding of video. *Child Development,* **69**, 950–965.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard Univ. Press.

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

Whitehurst, G. J., Arnold, D. H., Epstein, J. N., Angell, A. L., Smith, M., & Fischel, J. E. (1994). A picture book reading intervention in daycare and home for children from low-income families. *Developmental Psychology,* **30**, 679–689.

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.

Zaitchik, D. (1990). When representations conflict with reality: The preschooler's problem with false beliefs and "false" photographs. *Cognition,* **35**, 41–68.

Zelazo, P. D., Frye, D. & Rapus, T. (1996). An age-related dissociation between knowing rules and using them. *Cognitive Development*, **11**, 37–63.

Zelazo, P. D., Sommerville, J. A., & Nichols, S. (1999). Age-related changes in children's use of external representations. *Developmental Psychology,* **35**, 1059–1071.

Received October 30, 1997; revised January 19, 2000