

## Draw It Again Sam: The Effect of Drawing on Children's Suggestibility and Source Monitoring Ability

Maggie Bruck

*Johns Hopkins University*

Laura Melnyk

*McGill University, Quebec, Canada*

and

Stephen J. Ceci

*Cornell University*

Preschool children (aged 3 to 6 years) participated in a magic show. Later, the children were given repeated true and false reminders about the show. Half the children were asked to draw these true and false reminders (drawing condition) and half the children were asked questions about the reminders but not to draw them (question condition). Later, children in the drawing condition had better recall of true reminders than children in the question group; however, children in the drawing group also recalled more false reminders than children in the question group. Finally, although children in the drawing group had better memory of the source of the reminders than children in the question group, both groups equally reported that the false reminders actually happened. © 2000 Academic Press

*Key Words:* suggestibility; rehearsal; memory; source monitoring; children; drawing.

Obtaining complete and accurate reports of past events from young children is problematic for several interrelated reasons. First, when young children are asked open-ended questions, or asked to provide a free narrative of an event, their responses tend to be sparse. Although they can be quite accurate when there is

This research was carried out at McGill University when the first author was on academic staff. Portions of this research were supported by grants from the Natural Sciences and Engineering Research Council of Canada (to M. Bruck) and the National Science Foundation BSR-9312202 (to S. J. Ceci).

Address correspondence and reprint requests to Maggie Bruck, Department of Psychiatry and Behavioral Science, Division of Child and Adolescent Psychiatry, 600 N. Wolfe Street, Baltimore, MD 21287-3325.

no source of memory contamination (e.g., suggestive questions), and when the delays between the event and their report of it are not long, young children's free recall is often limited and incomplete (Baker-Ward, Gordon, Ornstein, Larus, & Clubb, 1993; Peterson & Bell, 1996; Steward & Steward, 1996). As a result, in order to elicit more complete information from children, interviewers often resort to a number of different conversational or interrogative strategies (e.g., see Ceci & Bruck, 1995). A typical strategy is to ask many specific questions including those that require only one-word responses and/or which presuppose specific answers (Warren, Woodall, Hunt, & Perry, 1996). These questions often can be classified as leading or misleading. When children do not provide an appropriate or anticipated answer, questions often are repeated (Poole & Lamb, 1998).

Although these interviewing techniques may elicit information that might not otherwise be reported, their use with young children presents difficulties. This is because young children's answers to specific questions, such as simple yes/no or choice questions, are quite inaccurate (e.g., Brady, Poole, Warren, & Jones, 1999; Peterson & Bell, 1996; Peterson & Biggs, 1997; Walker, Lunning, & Eilts, 1996). When questions are explicitly leading and misleading, preschoolers are more often led by such questions compared to older children (Ceci & Bruck, 1993; McAuliff, Kovera, & Viswesvaran, 1998). Finally, young children tend to change their answers in response to repeated specific questions (e.g., Poole & White, 1991), a strategy that can render their reports unreliable.

Because of these problems in obtaining complete and accurate information from young children while simultaneously avoiding sources of unreliability, researchers have attempted to identify nonsuggestive interviewing techniques. For example, the degree to which nonverbal cues, such as props and anatomically correct dolls, influence the accuracy of children's reports has been examined. It has been hypothesized that nonverbal cues might provide young children with the appropriate cues to facilitate memory retrieval as well as help them describe complex actions. The results of a number of studies, however, indicate that nonverbal props and cues often do not increase accuracy and completeness of young children's reports. Rather, although props increase the number of accurate details recalled, these also increase the number of inaccurate details provided (Gee & Pipe, 1995; Salmon, Bidrose, & Pipe, 1995; Salmon & Pipe, 1997; Steward & Steward, 1996; Wilson & Pipe, 1989). Similarly, studies with anatomically detailed dolls reveal that dolls generally do not affect the quality or quantity of young children's reports (e.g., Goodman & Aman, 1990; Saywitz, Goodman, Nicholas, & Moan, 1991), although they may encourage poor interviewing practices (Lamb, Hershkowitz, Sternberg, Boat, & Everson, 1996). In addition, some researchers have reported that dolls may promote more inaccurate reports, particularly in children under the age of 5 years (e.g., Bruck, Ceci, & Francoeur, 2000; Bruck, Ceci, Francoeur, & Renick, 1995; DeLoache, Anderson, & Smith, 1995; DeLoache & Marzolf, 1995).

Another technique that holds some potential for obtaining accurate and com-

plete reports from young children involves the use of drawing. This may provide a venue for children to spontaneously talk about what they are drawing, and the drawing itself may provide relevant cues to aid memory retrieval. Despite the common use of drawing in assessment and therapy with children (e.g., Burgess & Hartman, 1993; Burgess, McCausland, & Wolbert, 1981; Oppawsky, 1991; Rae, 1991; White, 1990), until recently there were no data on the effectiveness of this technique. Specifically, does drawing increase children's recall without simultaneously increasing errors? Most forensic investigators and therapists who use drawing to facilitate young children's understanding and recall of traumatic events assume that it does.

Three recent studies provide the first evidence that drawing may promote accurate verbal recall. In the first of these studies, Butler, Gross, and Hayne (1995) asked children open-ended as well as specific questions about a recent class trip to a fire station. One group of children drew their answers as they gave a contemporaneous verbal narrative; a control group did not draw but only gave a verbal narrative. For 5- and 6-year-old children, drawing promoted more complete reports without reducing accuracy. For 3- and 4-year-old children, however, drawing had no influence on the accuracy or completeness of reports.

In a second study, Gross and Hayne (1998) found that drawing facilitated young children's ability to talk about emotional experiences. Children produced more details about emotional events (e.g., "Tell me about a time you felt happy") when they could draw and tell compared to when they just could tell. Thus, asking children to verbalize what they are drawing may be a promising nonsuggestive interviewing technique for kindergarten children.

Finally, in an unpublished study, Brennan and Fisher (1998, reported in Poole & Lamb, 1998) found that children (ages 9 and 10) recalled significantly more details about a first-aid lecture when they were allowed to draw compared to children who did not draw. Accuracy rates were comparable for the two groups.

Although these three studies were not designed to reveal the basis of the drawing-memory facilitation effect, they indirectly shed light on this issue. Butler et al. (1995) proposed that drawing was effective because the product of the drawing allows children to provide their own retrieval cues that facilitate recall. These researchers noted that children in the drawing condition not only verbally described most of the items that they drew, but they also reported details that they did not draw. For example, in their fire station study, the children in the verbal control group typically responded to the question "How did you get there?" with "By bus," whereas children in the drawing group drew a bus and in their verbalizations about their drawing tended to provide more information, elaborating on the color of the bus, where the child was seated, who he or she sat with, and so forth. Thus, drawing may facilitate a child's recall because it allows the child to determine the salient aspects of the event, which in turn cues recall for other items. In this way, drawing may help to minimize or eliminate the

suggestive influence of the interviewer because the cues are largely child-generated rather than externally supplied.

In the present study, we extended work on the influence of drawing in three ways. First, whereas Butler et al. (1995) established that drawing *during the act of recall* facilitates retrieval of accurate information that otherwise might not be reported, the present study focused on the effectiveness of drawing when it is undertaken *prior* to the act of recall—as a rehearsal device. Specifically, in the present study children were given verbal reminders of an experienced event, and then children either drew these reminders or else answered questions about these reminders. Our goal was to determine whether drawing or answering questions was a better rehearsal strategy as determined by the accuracy and completeness of children's later recall. We had an avowedly practical reason for asking whether drawing long before recall was an effective memory aid: Although there are no data on the frequency that children are asked to draw in forensic and therapeutic contexts, it is known that they are sometimes asked to draw depictions of their presumed victimization in the course of investigations or therapy (e.g., Burgess & Hartman, 1993; Poole & Lamb, 1998, Chap. 6). When children are asked to draw in these contexts it is unclear what effect this may have on the accuracy of their later recall. Drawing a crime scene many weeks or months before being asked to recall it may be different from asking children to recall online during the act of drawing. Recently, Salmon and Pipe (2000) reported that drawing had no long-term facilitation on memory when it was carried out long before recall. So perhaps drawing facilitates online recall as the child is drawing, but after a sufficient delay between drawing and recall, the child's recall will not have benefited from having done the drawing—or even have suffered from it.

Second, whereas previous researchers focused on the potential benefits of drawing to facilitate memory, we examined the possible risks as well (see Bruck & Ceci, 1996, for a discussion of possible risks and benefits). Specifically, in previous studies children were asked to draw their answers to open-ended, nonleading questions dealing with actual (as opposed to fictitious) events. In the real world, however, interviewers do not always know the details of the event in question or even whether it actually occurred. Consequently, they may unwittingly mislead children or inadvertently provide inaccurate information during investigatory and therapeutic interviews. The issue that arises, therefore, is the degree to which being asked to draw misinformation that was inadvertently supplied by an interviewer damages the accuracy of children's subsequent reports.

For example, in the Canadian case of *R. v. Marquard*, a child who originally claimed to have burned her face with a cigarette lighter eventually told disbelieving interviewers that her grandmother had put her face on a hot stove. The grandmother was convicted of aggravated assault but later acquitted by the appeals court. A review of the record reveals yet another way in which drawing misinformation can later get incorporated into memory:

Q “When she put your face on her stove, what were her actions?”

A "She turned the controls on."

Q "How did the stove turn on? Do you know what a control does or looks like? Draw it as best you can and show me where the controls are."

(Child draws picture.)

Q "Is there a door on this stove with a window and handle?"

(Child adds features to her drawing.)

Q "Perfect."

Perhaps at the time the drawing was first made the child distinguished between details in the drawing that actually occurred and those that did not, but after a long delay the two may have become indistinguishable. To examine this issue, we asked children to draw true reminders as well as false reminders of an event they actually experienced. Many weeks later they were asked to recall the event, and their memory performance was compared to the performance of children who had also been given true and false reminders but not required to draw.

We hypothesized that drawing would be a good rehearsal device for true events because it adds to the purely verbal components of a verbal interview and may result in dual encoding of the information (Paivio, 1968): Children who draw the true reminders expend motor and mental effort, and spend more time thinking about each reminder than they might in a verbal interview alone. The extra effort required by drawing might aid recall by "tagging" the suggested events in memory and setting these apart as "special," linking them to their verbal codes, thus enhancing their later retrievability. In addition, the visual representation produced by drawing may help to establish the event in memory. This hypothesis is suggested by the findings of Foley, Aman, and Gutch (1987), showing that children could better discriminate pictures they traced with a pencil than those they traced with their finger (i.e., producing no visual representation on the paper). Foley et al. suggested that the kinesthetic feedback in conjunction with the external markers from visible consequences of actions increase memorability.

When children are asked to draw a false event, however, the extra time and effort required to draw may lead to more time "thinking about" or imagining the false event and this could increase later acceptance of the false information (e.g., Hyman & Pentland, 1996). Previously, Ceci, Loftus, Leichtman, and Bruck (1994) found that asking children to repeatedly think about a fictitious event and to elaborate its details in their mind led to increased acceptance of its reality. And recently Ackil and Zaragoza (1998) demonstrated that when young children are forced to confabulate (answer questions that are known to be false) the act of doing so leads to later memory errors. This type of memory error is evidenced even though the children can correctly tell at the time they are answering the question that it was false. So, drawing a false event, even if the children realize that it is false, may also lead to later recollective errors. This led to the third way the present study went beyond the existing literature on drawing.

The third way the present research extended prior work was to examine

whether requiring children to draw fictitious details proves more deleterious to the generation of false beliefs than requiring them to provide only verbal answers to questions based on the same fictitious details. The answer to this question may be related to children's source attributions of events. In the present study, source attribution refers to the children's ability to correctly determine if the events they are recalling had formerly been drawn or answered as well as if they actually occurred (Johnson, Hashtroudi, & Lindsay, 1993). This is an important issue in the developmental literature because it has been found that young children make a disproportionate number of source misattributions; that is, often they have difficulty distinguishing events that they actually experienced from those they were merely told about (e.g., Ackil & Zaragoza, 1995; Parker, 1995; Poole & Lindsay, 1995, 1999).

Recent work by Poole and Lindsay (1995, 1999) demonstrates how preschoolers' susceptibility to source monitoring errors may underlie their suggestibility. In their initial study (Poole & Lindsay, 1995), preschoolers played with "Mr. Science" for 16 min in a university laboratory. During this time, the child participated in four demonstrations (e.g., lifting cans with pulleys). Three months later, the children's parents were mailed a storybook, specially constructed for each child, that contained a biographical description of their child's visit to Mr. Science. However, not all of the information was accurate: Although the story described two of the experiments that the child had seen, it also described two that the child had not seen. The parents read the story to their children three times. When later interviewed by the experimenters, the children reported that they had participated in demonstrations that, in actuality, had only been mentioned in the stories read to them by their parents. When asked to source monitor their assents ("Did Mr. Science actually do that?" and "Did your Mom read you a story about that?"), most children said that the false events really had happened. Thus, the children in this study had particular difficulty identifying the source of a suggestion; they confused their parents' reading the suggestion to them with their experiencing the act entailed in the suggestion. Poole and Lindsay (1999) recently replicated these findings with children from a wider age range (3 to 8 years old). The findings were similar across ages with one exception: The source monitoring procedures enabled the older but not the younger children to reduce the rate at which they falsely reported having experienced the suggested events. When asked questions such as "Did Mr. Science really do that with you?" and "Did your Mom read you this in a story?", the older children were more likely to recant their previous claims and assert that their mother had told them.

The design of the present study allows a further examination of the source monitoring skills of young children. There are two opposing predictions concerning the degree to which drawing may impair or facilitate source monitoring. First, drawing may impair source monitoring judgments about false reminders. That is, children who draw are more likely to claim that the false suggestions actually happened compared to children who do not draw. This prediction is

based on previous findings of Hyman and colleagues (Hyman, Gilstrap, Decker, & Wilkinson, 1998; Hyman & Pentand, 1996) that explicit instructions to use mental imagery increase the likelihood of adults' creating false memories that reflect source monitoring errors. Because drawing results in the production of visual images, one would predict that it also results in the loss of the source of the false reminders.

An alternative plausible hypothesis, about which there are few direct data, concerns the relative distinctiveness or explicitness of the two sources of reminders. We suggest that the more memorable or explicit the context of a suggestion, the more likely its source will be retained. If the suggestion, however, is couched in normal activities and not brought to the child's attention, then its source may be more likely to be forgotten for later recall. Foley and colleagues (Foley et al., 1987; Foley, Durso, Wilder, & Friedman, 1991) found that implicit compared to explicit instructions to form images of objects resulted in more source monitoring errors for both children and adults. In the present study, drawing was more memorable or explicit than answering questions about reminders. For the drawing condition, the child must pay attention to the reminder and draw it. For the question condition, the child hears the reminder one time, and then answers a question about it. It may be more difficult for children to remember what was said within the context of an ordinary interview than in the context of one where they have to draw the answers.

## METHOD

### Design

Preschool children individually interacted with a magician who performed two tricks. Each child was then interviewed on three separate occasions about the magic show. In the first interview, the children were given true and false reminders of details in the magic show. Half the children were told to draw the reminders (the drawing condition) and the other half answered questions about the reminders (the question condition). In the second interview, the same procedures were repeated. In the exit interview, children's recall of the details of the magic show and their source attributions about these recalled details were assessed.

### Participants

Eighty-seven English-speaking children attending preschool or kindergarten programs were recruited from four schools in the Montreal area. The children were between the ages of 44 and 75 months and were predominantly from middle-class socioeconomic backgrounds.

Children were randomly assigned to a drawing group ( $n = 45$ ) or to a question group ( $n = 42$ ). The two groups were equated in terms of gender (56% and 55% of the children were girls in the drawing and question groups, respectively) and age ( $M = 65$  months for both groups).

## Procedure

### *Session 1: Magic Show*

A female experimenter visited the classroom and took one child at a time to another room in the school. To make the child comfortable, she engaged the child in conversation while playing with a sticker book. After a few minutes, the experimenter then performed a magic show that contained the following events (the italicized portions represent 16 prespecified target events about which children were later reminded and questioned):

The *magician put on a black hat* and asked the *child to put on a red magic helper cape*. The magician did two tricks. In the first trick, *the magician put a red ball in her pocket*. The *magician waved her wand* and then the ball reappeared in a small blue container. In the second trick, *the child poured water into a cup* and then *poured the water from the cup into a bowl*. In another magic cup, *the magician stirred water with a spoon* so that when the child tipped this magic cup over, nothing came out.<sup>1</sup> *The magician put the magic cup in the corner* and said, "Stay, magic cup, stay!" After the trick, *the magician tripped over her shoelaces* and *laughed as she fell down*, saying how clumsy and silly she was. The magician said that *she hurt her hand*, and *she put a bandage on her hand*. At clean-up time, *the magician put the tricks away in a box*, and *brushed her magic stage* with a little brush. *She showed the child a picture of her new puppy*, and *gave the child an award* for being such a good helper.

### *Session 2: First Suggestive Interview*

Approximately 2 weeks after the magic show ( $M = 16$  days,  $SD = 2$ ), a female experimenter, who was not the magician, interviewed the child. There were three parts to this interview: warm-up, free recall, and suggestion.

#### Warm-up

The interviewer began by telling children in the *drawing condition* that she collects children's drawings and that she likes to show them to her friends and teachers. The children were then asked to draw their brothers and sisters as well as a school activity.

The interviewer told children in the *question condition* that she collects stories that children tell her and that she likes to share these with her friends and teachers. The interviewer told children in the question condition that she would record their stories on her tape recorder. At this point, the experimenter turned on her tape recorder and asked the children about their brothers, sisters, and school activities.

<sup>1</sup> A powder called "Lightning Gel" was mixed into the water, causing it to congeal on contact.



### Baseline Free Recall

The experimenter asked the child to draw (or in the question condition, to tell) her two things that happened with the magician.<sup>2</sup> The experimenter prompted the second item by saying, "That's really neat; could you draw/tell me another thing that happened with the magician?"

*Scoring.* The number of utterances produced in the child's free recall was counted. An utterance is a statement, bounded by pauses, that contains one verb. For example, "The magician wore a black hat and I wore a magic cape" contains two utterances: (1) The magician wore a black hat, and (2) I wore a magic cape. Utterances were classified as accurate or inaccurate.

### Suggestion Phase

The magician event contained 16 prespecified target events, which were divided into four scenes (ball trick, water trick, fall, and clean-up), with four events in each scene. As shown in the Appendix, one false event was generated from each true target event. For example, the target event "The magician hurt her hand" generated the false event "The magician hurt her leg."

Each child was given four true reminders (actual target events) and four false reminders (generated false events). These were selected in the following way. One true reminder and one false reminder were taken from each of the four scenes. True reminders and the paired generated false reminder were never presented together. Thus the same child would never receive the true reminder "The magician hurt her hand" with the false reminder "The magician hurt her leg." (As described below, the remaining nonreminded eight target events were used as control items on the recognition test.) In order to ensure that all items were equally used and that different combinations of true and false reminders were used across participants, 16 different interview schedules were randomly selected from a large pool of possible choices. Assignment of interview versions was counterbalanced across conditions so that the same interview schedules were equally used in the draw and question conditions.

In the drawing condition, the interviewer began the suggestion phase by saying that she had heard about some things that happened with the magician, and maybe the child could draw these things for her. For example, the experimenter said, "I heard that the magician fell and hurt her leg. I want you to draw that for me." If a child resisted a reminder (e.g., the child said that a reminder did not happen), the experimenter repeated the reminder, saying, "Well, let's just pretend that [the reminder] happened; could you draw that for me?" The remaining true and false reminders were randomly presented in this same manner. The experimenter made comments throughout the interview, such as "You're a good

<sup>2</sup> In pilot testing, we found that some children would take up to 20 min to draw a picture when there were no time constraints. Because of time constraints, the final procedures involved asking the child to recall or to draw two items.

drawer." After all the items were drawn, the interviewer repeated the reminders, preceded by "You remembered that. . ." or "You drew that. . ."

In the question condition, the experimenter told each child that she had heard certain things about the magician's visit. She then told the child one of the reminders (e.g., "I heard that the magician gave you a hug") and then asked the child a question that was related to the event (e.g., "Did the magician give a hug to any other kids in the class?"). When a child resisted answering a question, the experimenter repeated the reminder and the question and asked the child to pretend that the reminder had occurred. The remaining true and false reminders were randomly presented in the same manner. As was the case in the drawing condition, the experimenter made comments throughout the verbal interview, such as "You're a good storyteller." Once all the items had been presented, the interviewer repeated each of the items in the suggestion phase, randomly preceded by "You remembered that. . ." or "You told me that. . ." For reasons presented below, children in the question group played with Lego blocks between the reminder items throughout their interview.

Pilot testing indicated that interviews in the drawing condition were longer than those in the question condition. In order to equate the interview length across the two groups, the children were tested in blocks. A group of drawing children were tested first, and the length of each interview was measured. Then a similar number of children were tested in the question condition, each paired with a child in the drawing condition. The interviewer allowed the child in the question condition to play with Lego blocks between the reminder items so that the question interview approximated the length of the matched drawing interview. Although this equated the total time spent with the interviewer in both conditions, drawing by its nature entails more time than answering questions, resulting in an inherent difference in the amount of time spent on the actual task.

### *Session 3: Second Suggestive Interview*

Approximately 12 days after the first suggestive interview ( $M = 12$  days,  $SD = 3$ ), the same experimenter returned for a second suggestive interview. The experimenter told each child in the drawing condition that the pictures were ruined because they had fallen out of her bag. She told the children in the question condition that the tape with their stories was ruined because it had fallen out of her bag. In both conditions, she told the children that she really wanted to share with her teachers and friends the children's pictures or stories, so she came back to get these pictures or stories again.

The second interview followed the same procedures as the first interview, except there was less free play and no free recall; as a result the total average time of the second suggestive interview ( $M = 8.4$  min,  $SD = 3.6$ ) was shorter ( $p < .01$ ) than the average time for the first suggestive interviews ( $M = 13.6$  min,  $SD = 4.9$ ). The same items were suggested and rejections of the false reminders

were dealt with in the same manner as in the first interview.<sup>3</sup> As before, the testing was conducted in blocks, with question children matched with their drawing counterparts to equate the interview lengths.

#### *Session 4: Exit Interview*

Approximately 2 weeks after the second suggestive interview ( $M = 12$  days,  $SD = 3$ ), the exit interview was conducted by a new female experimenter. Children's recall of the events of the magic show was assessed by free recall questions and by a recognition test. Source monitoring for all target items reported in the recall tasks was assessed by free recall questions and a recognition test. Responses to source monitoring free recall questions are termed "spontaneous source monitoring" and responses to source monitoring recognition items are termed "probed source monitoring."

#### Recall of Magic Show: Free Recall

The interviewer told the child that she was not present for the magician's visit. She then asked the child to tell her everything that had happened during that visit. After the child gave a response, the interviewer asked, "Can you tell me one more thing?" This continued until the child could not report any more events.

*Scoring.* The number of utterances in the free recall was counted. Each utterance was categorized in terms of accuracy and also whether it had been reminded during the suggestion phase of the study. Utterances not relevant to the activities of the magic show were excluded (e.g., "It was fun"; "I liked the magician").

#### Recall of Magic Show: Recognition Test

The child was asked 16 yes/no questions. Four questions probed the true reminders that were given to the child ("Did the magician hurt her hand?") and four probed the false reminders. In addition the child was asked four questions about true nonreminded events (true control items) and four questions about false nonreminded events (false control items) that had not been used as reminders in the suggestive interviews for that child. This combination of true and false control items was dictated by the composition of the specific interview schedule that was assigned to each child, creating 16 different exit interview schedules. As was the case for the reminders, a true control item and a false control item were selected from each of the four scenes. Each control item could not be from the same pair as any other item in the recognition test. Thus for each scene there was exactly one question about a true reminder, a false reminder, a true control, and a false control item.

The experimenter gave each child two practice questions to make sure the

<sup>3</sup> Although there was a significant decrease in rejections of false reminders from the first interview (15%) to the second interview (7%),  $F(1, 85) = 11.83$ ,  $p < .01$ , there were no main effects or interactions involving the between-group factor of condition (i.e., draw vs question).

child understood the task. If the child did not seem to understand, the experimenter repeated the instructions, and gave the child two more practice questions.

If any of the events in a child's recognition interview version were reported in free recall, the child was automatically credited with a "yes" response and this item was not asked in the recognition task.<sup>4</sup> For example, if the child's free recall included the false reminder "The magician dropped her wand," the question "Did the magician drop her wand?" was not asked; nonetheless the child was credited with a "yes" response for this item.

*Scoring.* Accuracy was assessed by counting the number of assents to true reminder and true control items and the number of denials to false reminder and false control items.

### Spontaneous Source Monitoring

The experimenter reminded the child that a lady visitor asked the child to draw some things about the magician's visit (or to use the tape recorder to talk about some things about the magician's visit). The experimenter told the child that she was not present when the "drawing lady" or the "story lady" came, and then she asked, "What did she tell you to draw (or tell you) that *really happened* when the magician came?" The experimenter prompted for more items until the child could not report any more.

*Scoring.* True reminder items were counted as accurate responses.

The experimenter then asked if the drawing (or story) lady made any *mistakes* about what she told the child to draw (or whether she made any mistakes in what she told the child) about the magician's visit. If the child assented, the experimenter asked, "What mistakes did she make? What did she tell you to draw (tell you about) that didn't really happen when the magician came to do tricks?" The experimenter again prompted for more responses until the child could not report any further details.

*Scoring.* False reminder items were counted as accurate responses.

### Probed Source Monitoring

Two questions were asked about each item that had been reported in the recall tasks: (1) "Did it really happen?" and (2) "Did the drawing lady (story lady) tell you to draw it (tell you about it)?" For example, if the child previously recalled that the magician gave her a hug, then she was later asked, "Did the magician really hug you?" and "Did the drawing lady tell you to draw a picture of the magician hugging you (Did the story lady tell you that the magician hugged you)?" If the child had provided a response about an event in the spontaneous source monitoring, then this event was not probed; nevertheless the questions

<sup>4</sup> The "Reminded" columns in Table 1 show the mean number of reminded items generated in free recall. Each of these means subtracted from 4 yields the number of questions asked on the recognition task for the true reminder and false reminder items. Because children rarely provided free recall utterances of the content contained in the control recognition questions, for most children there were four false control and four true control questions asked on the recognition test.

were scored for the spontaneous statement as though the content of the spontaneous utterances had been probed. For example, if the child spontaneously generated the false reminder statement "The magician hugged me" in response to the question "What did she tell you to draw (or tell you) that really happened when the magician came to do tricks?", a "yes" response was recorded for the two probe questions (i.e., "Did that really happen?" and "Did the lady tell you that?") that would have been asked.

*Scoring.* Accurate responses for the "Did it really happen?" question were assents for true reminder and true control items and denials for false reminder and false control items.

Accurate responses for the "Did the lady tell you?" question were assents for the true reminder and false reminder items and denials for the true control and false control items.

## RESULTS

Analyses were carried out on (1) recall of the magic show prior to reminders, (2) recall of the magic show in the exit interview after reminders, and (3) source monitoring of all target items reported in the exit interview recall tasks. For all analyses the between-group factor was condition (drawing vs question). Initial analyses examined the effects of age on performance. For all analyses reported, there were no significant main effects or interactions involving age and therefore this factor was not included in the model.

### Baseline Measure: Free Recall in First Suggestive Interview

The free recall data were analyzed to determine if the children in the drawing and question conditions differed with respect to their free recall accuracy before suggestions were provided.

A  $2 \times 2$  ANOVA with repeated measures was carried out on the total number of utterances. The independent variables were condition (drawing vs question) as the between-group factor and utterance accuracy (accurate vs inaccurate) as the repeated factor. There was a main effect of accuracy,  $F(1, 85) = 100.13, p < .001$ . There was no main effect or interaction involving the factor of condition. The children provided more accurate utterances ( $M = 3.5, SD = 3.0$ ) than inaccurate utterances ( $M = .1, SD = .3$ ). In fact, there were very few inaccurate utterances: only nine children (10.3% of the sample) had an error in their prereminder free recall.<sup>5</sup> This indicates that before any suggestive interviewing occurred, free recall was similar across experimental conditions and the children's recall of the event was extremely accurate.

<sup>5</sup> Because of these floor effects, the analysis was carried out on square root transformations of the data. The results were identical to those reported in the text.

TABLE 1  
Mean Free Recall Utterances in Exit Interview

	Condition			
	Draw		Question	
	Reminded	Not reminded	Reminded	Not reminded
Accurate	1.5 (1.0)	3.5 (3.6)	1.4 (0.9)	4.5 (3.7)
Inaccurate	1.2 (1.1)	.3 (0.6)	.8 (0.8)	.5 (1.1)

*Note.* Standard deviations are in parentheses.

### Exit Interview: Recall of Magic Show

#### *Free Recall*

The number of utterances in free recall served as the dependent variable in a 2 (condition)  $\times$  2 (accuracy)  $\times$  2 (reminded vs not reminded) analysis of variance with repeated measures on the last two factors. There were main effects of accuracy,  $F(1, 85) = 85.68$ ,  $p < .01$ , and reminder status,  $F(1, 85) = 18.38$ ,  $p < .01$ .<sup>6</sup> These main effects were qualified by two significant interactions: Condition  $\times$  Reminder,  $F(1, 85) = 3.91$ ,  $p < .05$ , and Accuracy  $\times$  Reminder,  $F(1, 85) = 73.81$ ,  $p < .01$  (see Table 1).

Although there were no main effects of condition, planned comparisons of the Condition  $\times$  Reminder interaction reveal that a larger proportion of the utterances in the drawing condition incorporated the reminders (50%) than was found in the question condition (34%),  $p < .05$ .

The Accuracy  $\times$  Reminder interaction reflects the findings that although there were similar numbers of true and false reminders incorporated into accurate recall (1.5 vs 1.0), most of the inaccurate utterances contained false reminders (52%), whereas most of the accurate utterances contained nonreminded items rather than true reminders (i.e., only 36% of accurate utterances contained a true reminder item),  $p < .01$ .<sup>7</sup>

Finally, in all comparisons for the significant interaction terms, children produced more accurate ( $M = 5.6$ ,  $SD = 4.0$ ) than inaccurate ( $M = 1.4$ ,  $SD = 1.3$ ) utterances.

To summarize, although interviewing condition did not influence the produc-

<sup>6</sup> Children reported false reminders that conflicted with the actual events at the same rate as they reported false reminders that did not conflict with the actual events.

<sup>7</sup> Accurate utterances contained many more nonreminded than reminded items because there were only four accurate reminders that the child could report whereas there were a large pool of nonreminded items that could be reported.

TABLE 2  
Total Recall Accuracy Mean Scores

	Condition							
	Draw				Question			
	True reminder	False reminder	True control	False control	True reminder	False reminder	True control	False control
Raw score	3.83 (0.44)	.78 (1.15)	2.44 (1.21)	3.11 (1.14)	3.44 (0.74)	1.44 (1.27)	2.54 (0.93)	3.32 (0.99)

*Note.* The maximum for each cell is 4. Standard deviations are in parentheses.

tion rate or the accuracy rate of children's free recall, children in the drawing condition were more likely to include true and false recalls in their free recall than children in the question condition.

### *Total Recall*

Accuracy of total recall was evaluated by summing accurate utterances about target events produced in free recall and accurate responses on the recognition test (see Table 2).

The total number of accurate responses were subjected to a three-way repeated ANOVA with condition (drawing vs question) as the between-subject factor and with item accuracy (true items vs false items) and item type (reminder items vs control items) as the repeated factors.

The analysis yielded significant main effects of reminder,  $F(1, 85) = 31.43$ ,  $p < .001$ , and accuracy,  $F(1, 85) = 42.68$ ,  $p < .001$ . In addition, the following two-way interactions were significant: Accuracy  $\times$  Condition,  $F(1, 85) = 4.27$ ,  $p < .05$ , and Item Type  $\times$  Accuracy,  $F(1, 85) = 268.91$ ,  $p < .001$ . There was also a significant three-way interaction of Item Type  $\times$  Accuracy  $\times$  Condition,  $F(1, 85) = 6.52$ ,  $p < .01$ .

The three-way interaction was examined using planned comparisons. (Although raw scores were used in these analyses, for clarity the results are presented in terms of percentage correct—the raw scores were divided 4.) For the control items (i.e., items that had not been used as reminders in the suggestive interviews), the children in the drawing and the question conditions were equally accurate in correct assents to true control items (drawing = 61%, question = 63%) and in correct denials to false control items (drawing = 77%, question = 83%). However, the two groups did differ on the reminded items: The drawing group children were more accurate on true items (96%) than the question group children (86%). For false reminders, the opposite pattern was obtained: The drawing group children were less accurate in denying these items (correct denials = 20%) than were the question group children (correct denials = 36%).

The three-way interaction was obtained because of the relatively poor performance of the drawing children on the false reminders. Between-group comparisons for false reminders were significantly larger than between-group comparisons for true reminders ( $p < .01$ ).

Finally, it is important to note that although there were group differences on the reminded items, both groups of children showed sizable and significant misinformation effects (accuracy rejecting false reminders was lower than accuracy rejecting false controls) and sizable facilitation effects (accuracy on true reminders was higher than accuracy on true controls).

Thus, drawing specifically mentioned information influences children's reports of a past event. Although drawing increases recall of true reminders, it is coupled with increased acceptance of false reminders. As shown in Table 2, the size of this effect may be constrained by the ceiling effect obtained for the true reminder items, while the effect is especially pronounced for the false reminder items.

#### Post Hoc Analyses: Factors Associated with the Drawing Effect

The finding that drawing had both facilitative and deleterious effects on recall could be due to a number of factors that are not specific to drawing. For example, it is possible that attentional focus per se could account for the between-group differences. That is, even though the two groups were equated for time on task, the drawing task engaged the child more fully than the nondrawing task where children filled in the time by playing with Lego blocks between the reminders. The increased rehearsal and misinformation effects associated with drawing could also reflect factors more intrinsic to drawing, namely that having a pictorial representation of the event influenced recall more than having a linguistic reminder of the event. In considering these two (not mutually exclusive) explanations of the drawing effects, there is no obvious way to make the nondrawing condition comparable to the drawing condition in terms of attentional demands without engaging the child in the nondrawing condition in more verbalizations. Although this might increase attentional focus, the effects on recall would be confounded by the fact that the repetitions of suggestions by themselves influence recall. Similarly it is difficult to control for attentional focus in the drawing condition because by its very nature drawing requires attentional focus.

Nevertheless, in order to explore these issues further we carried out the following analyses for children in the drawing group. First we correlated the accuracy of their recall responses with measures of the quality of their drawing. If it was pictorial representation that accounted for the effects of drawing, then children whose pictures were more detailed should have better recall of true events and greater misinformation effects.

Each drawing was rated for the degree of pictorial representation on a 5-point scale. A coding scheme for each of the 16 true and 16 false items was devised. For example, for the false reminder "You wore red magic helper boots," the following categories were devised from most to least representative: a child



wearing red boots, a child wearing two nonred boots, two red boots, one red boot, two nonred boots, one nonred boot, squiggles or lines that had no representation. All the drawings of "red boots" reminder were then ranked in terms of these criteria. Interrater reliability was high overall (97% agreement across all judgments). The child's average rating for drawings of false reminders was correlated with recall scores on these items. Similarly the child's average rating for drawings of true reminder was correlated with recall on these items. None of the correlations neared significance.

Next, in order to evaluate attentional focus, we correlated recall with the amount of time that it took each child to draw the true reminders and the false reminders. We reasoned that time on task was a proxy measure of attentional focus. There were no significant correlations between drawing time of false items and true items with respective recall of these items on the recall tasks.

Clearly these are only post hoc analyses and further work is required to examine the cognitive basis of the drawing effects. Although it was not an objective of the present study, at this time we have no data to explain the mechanisms underlying the drawing effects.

### Control Items versus Control Group

The finding that true reminders increased accurate recall and that false reminders increased recall errors is based on within-group comparisons of reminded vs control (nonreminded) items. Although this is a standard design to evaluate the effects of suggestion in most misinformation studies, and although there is some evidence that rehearsal affects only the rehearsed items (Cassidy & DeLoache, 1995), it could still be argued that a control group of children who receive no reminders between the magic and the exit interview is required to accurately evaluate the effects of true and false reminders. That is, it is possible that in the present study, performance on control items was influenced by the reminders during the suggestive interviews and therefore an uncontaminated baseline measure of control items is required. For this reason, we tested another 27 children of the same age ( $M = 65$  months, range 53–77 months) and neighborhood as the children in the drawing and question conditions.

Each of the 27 control children participated in a magic show exactly as described above but there were no suggestive interviews. Approximately 40 days after the magic show ( $M = 39$  days,  $SD = 4$ ), the exit interview was conducted by a new female experimenter. This delay was equated with the total delay between the magic show and the exit interview for the drawing group ( $M = 40$  days,  $SD = 4$ ) and the question group ( $M = 40$  days,  $SD = 5$ ). The control children's recall of the magic show was assessed by free recall and performance on a 16-item recognition test. There were no source monitoring questions in the exit interview because there were no reminders given to these children.

The free recall was carried out exactly as for the drawing and question groups. The recognition task differed in one respect—because the control children were

not given any reminders, there were no true reminder or false reminder questions. That is, all the questions were considered control items because they probed nonreminded events. Eight of the questions probed true events from the magic show, and eight of the questions focused on false events. The same interview schedules that were constructed for the drawing and nondrawing groups were given to the control children; thus the same questions were asked for all three groups.

For free recall, the control children provided as many correct details ( $M = 6.1$ ) as the drawing ( $M = 5.0$ ) and question groups ( $M = 5.9$ ),  $p < .40$ . However, the control children produced fewer errors ( $M = .37$ ) in free recall than did the drawing and verbal groups ( $M = 1.4$  and  $1.3$  errors),  $F(2, 111) = 6.79$ ,  $p < .01$ . This result reflects the fact that the control group was not given any misinformation between the magic show and the exit interview. If one excludes from the data all errors that directly reflect the false reminders provided to the drawing and question groups, then the control children do not differ from these groups on any measures.

Pooling the free recall and probed recognition data, the control group accurately recalled 70% of the true items. Their accuracy rate is the same as those for the two reminder groups' responses to true control (nonreminded) items (60% for drawing group and 63% for question group),  $p < .28$ . However, the 70% accuracy rate of the control group for the true items is significantly lower than the accuracy rates for the true reminder items for the drawing group (96%) and for the question group (86%),  $F(2, 111) = 18.92$ ,  $p < .01$ . Thus the beneficial effects of reminding are still observed when a nonreminded group is added to the analysis.

A similar pattern of results was obtained for the false items. Pooling the free recall and probed recognition data, the control group accurately denied 79% of the false items. This rate of accurate denials is the same as those for the two reminder groups' responses to the false control (nonreminded) items (77% for drawing group and 83% for question group),  $p < .61$ . However, the 70% accuracy rate of the control group for the false items is significantly higher than the rates of accurate denials for the false reminded items for the drawing group (20%) and for the question group (36%),  $F(2, 111) = 38.7$ ,  $p < .01$ . Thus the detrimental effects of providing false reminders are still observed when a nonreminded group is added to the analysis. Furthermore, it is interesting to note that for both the true and the false items the effects of reminding were specific to those items and did not generalize to the control nonreminded items. If there had been a general facilitation effect then the two reminder groups should have achieved higher accuracy scores on nonreminded true control items than the control group. This did not occur. Similarly, the two reminded groups had the same scores on the false control items as the control group. Thus providing misinformation in this study did not negatively affect memory for other details in the original event.

## Source Monitoring

In this section, we report the source monitoring data for items that had been assented to on the recognition and free recall tasks. The major objectives were to determine if the children could remember which items had been reminded and to determine whether the children would come to confuse false suggestions with actual events. Of specific interest was the degree to which loss of source (source misattributions) was greater for the drawing or question condition.

### *Spontaneous Source Monitoring Questions*

Children were asked two questions: (1) "What did she tell you to draw (or tell you) that *really* happened when the magician came?" and (2) "What did she tell you to draw (or tell you about) that *didn't* really happen when the magician came to do tricks?" These two questions generated 122 responses from the drawing group and 30 responses from the question group. Although the drawing group clearly provided more spontaneous responses than the question group ( $p < .01$ ), the relative accuracy of the responses were similar across groups: 57% of all responses made by the drawing group were accurate and 60% of all responses made by the question group were accurate ( $F < 1$ ). The major type of error made by both groups was providing false reminder items for the first question—that is, claiming that the false reminders really happened: This source monitoring error accounted for 36% of all the spontaneous responses of the drawing group and 30% of responses of the question group. The remaining errors involved providing true control items for the first question (7% drawing vs 10% question). This is an error because the true control items were not reminded.

### *Total Source Monitoring (Spontaneous Source Monitoring $\pm$ Probed Source Monitoring)*

In addition to generating spontaneous source monitoring statements, children also responded to two probed source monitoring questions for each particular item that had been assented to on the recall tasks and that was not generated in spontaneous source monitoring. For each such item, the child was first asked whether or not the item really happened ("Did [event] happen when the magician came?"). Second, the child was asked whether or not the lady told the child about the event. As explained in the method section, spontaneous source monitoring scores were added to the probed scores to yield a total source monitoring measure. Three separate analyses were conducted on these data: (1) correct HAPPENED attributions, (2) correct TOLD attributions, and (3) correct combined HAPPENED and TOLD source attributions.

### HAPPENED Attribution Accuracy

This analysis reflects the child's ability to accurately differentiate between falsely reminded events and true events. That is, it reflects the degree to which the false reminder comes to be believed as a real event. Furthermore, it serves as

TABLE 3  
Source Monitoring Responses for Happened and Told Questions: Percentage Correct

	Condition					
	Draw			Question		
	True reminder	False reminder	True control	True reminder	False reminder	True control
Happened <sup>a</sup>						
% Accurate	.98 (.07)	.17 (.25)	.92 (.24)	.97 (.09)	.10 (.24)	.97 (.11)
Told <sup>b</sup>						
% Accurate	.94 (.18)	.95 (.17)	.57 (.42)	.57 (.36)	.58 (.36)	.52 (.43)
Happened and Told <sup>b</sup>						
% Accurate	.91 (.19)	.16 (.27)	.53 (.43)	.54 (.37)	.04 (.11)	.54 (.43)

Note. Standard deviations are in parentheses.

<sup>a</sup>  $n = 45$  for drawing;  $n = 41$  for question.

<sup>b</sup>  $n = 41$  for drawing;  $n = 37$  for question.

a recantation measure (the degree to which assents to false reminders in the recall phase change to denials in the source monitoring phase). If the children came to believe that the false reminders actually happened, recantation rates should be relatively low.

Separate accuracy scores were calculated for the true reminder, false reminder, and true control items. The numerator for each score was the total number of correct responses to happen questions and the denominator was the number of items assented to on the recall tasks. For example, the relative accuracy for the true reminder items was total accurate source monitoring responses for true reminder happen probes divided by the number of true reminder items generated in the recall tasks. The data are shown in the top third of Table 3. False control items were not included in this and all other source monitoring analyses because these were rarely reported in spontaneous source monitoring (there were only two children in each condition who made such errors) and, as shown in Table 2, children rarely assented to these items on the recognition test.

The results of a 2 (condition)  $\times$  3 (item type) analysis of variance yielded a significant main effect of item type,  $F(2, 152) = 495.39$ ,  $p < .001$ .<sup>8</sup>

<sup>8</sup> All source monitoring analyses were rerun using arcsine transformations. The results were identical.

Collapsing across conditions, children were highly accurate on true reminder (98%) and true control (94%) items; they accurately reported that these events had actually happened when the magician visited. Accuracy rates on false reminder items were very low. That is, children in both conditions, for the most part, held onto their previous claims from the recall tests that the false reminder items had actually happened; only 14% of all previous false claims were recanted in this portion of the experiment and the rate of recantation was similar across conditions.

### TOLD Attribution Accuracy

This measure reflects the child's ability to accurately report which items were drawn (or told about). In other words, it is the ability to keep track of the reminded items. Separate scores were derived for true reminder, false reminder, and true control items. For example the relative accuracy for the false reminder items was the total accurate source monitoring responses for false reminder told questions divided by the number of false reminder items generated in the recall tasks.

A  $2 \times 3$  repeated ANOVA was conducted on the percentage of correct TOLD responses with condition (drawing vs question) as the between-group factor, and with item type (true reminder vs false reminder vs true control) as the repeated factor. There was a significant main effect of condition,  $F(1, 76)^9 = 44.54, p < .001$ , and a significant main effect of item type,  $F(2, 152) = 10.08, p < .001$ . There was a significant two-way interaction of Condition  $\times$  Item Type,  $F(2, 152) = 5.82, p < .01$ . The interaction is easily interpreted. Children in the drawing condition were more accurate than children in the question condition for true reminder and false reminder items, whereas there were no between-group differences for true control items. In other words, children in the drawing group were very accurate in remembering the items that they drew, whereas children in the question condition performed around chance: They could not reliably remember which items had been reminded. The drawing children's advantage is not simply due to a bias of replying "yes" to all questions. If this had been the case, one would have expected the children to continue with high rates of "yes" responses to true control questions, but this did not occur.

### Identification of the Sources of Individual Events: HAPPENED and TOLD Accuracy

This analysis reflects a combination of the previous two. Accuracy in this analysis reflects the child's ability to accurately report whether each event assented to in the recall tasks happened and was told. Pooling the spontaneous and probed source monitoring data, accurate responses for true reminder items

<sup>9</sup> Eight subjects were eliminated from this analysis because they did not assent either to any false reminder or to any true control items in the recognition test and thus could not be asked any source monitoring questions.

were "Yes it happened" and "Yes it was told." Accurate responses for false reminder items were "No it did not happen" and "Yes it was told." Accurate responses for true control items were "Yes it happened" and "No it was not told." Each of these scores was divided by the number of items in that category that were produced in the recall tasks (see bottom third of Table 3).

A  $2 \times 3$  repeated ANOVA was conducted on the percentage of correct responses with condition (drawing vs question) as the between-group factor, and with item type (true reminder vs false reminder vs true control) as the repeated factor. There was a significant main effect of condition,  $F(1, 76) = 20.33, p < .001$ , and a significant main effect of item type,  $F(2, 152) = 76.33, p < .001$ . There was a significant two-way interaction of Condition  $\times$  Item Type,  $F(2, 152) = 5.64, p < .01$ . Neuman-Keuls comparisons of the significant interaction indicated that the children in the drawing condition were more accurate at source monitoring the true reminder items than the children in the question condition. The two groups were equally accurate on the false reminder and true control items.

There are three important findings of the source monitoring analyses that will be elaborated in the discussion. First, children had better memories for reminders that they drew compared to reminders that they heard. Second, children in the drawing and question conditions rarely recanted their previous assertions that the false reminders had happened. Third, the drawing children's better memories for reminders in combination with their low recantation rates of previous false assents account for the findings in the last analysis that the drawing children have better identification of sources for true events but not for false events. These results highlight two separable elements of source monitoring skill and show that although drawing produces better source monitoring for the source of the suggestion, it does not protect children from confabulation (claiming that false events really happened).

## DISCUSSION

Previous work by Hayne and colleagues (Butler et al., 1995; Gross & Hayne, 1998) provides empirical support of the use of drawings to promote accurate and complete recall in preschool children. Our data support but also qualify these conclusions when drawing is used as a rehearsal device. First, as hypothesized, drawing does promote accurate recognition of true items. Second, drawing also promotes accurate source memory of reminded items. Children in the drawing group were better able than children in the question group to remember the source of the reminded items, whether they were for true or false events.

We also found, however, that there were certain negative consequences to using drawing as a rehearsal technique. First, drawing increased children's acceptance of false reminders. Children came to make false claims about a range of events; some of their false reports involved bodily contact ("The magician gave me a hug"), others involved misreporting of emotions ("The magician cried

when she fell"), and others involved salient actions in the event (e.g., "The magician hurt her leg"; "I put a Band-Aid on her"). Thus the deleterious effects of drawing on recall were not confined to peripheral events, but involved central actions, some of which involved the child. Second, drawing did not reduce confabulations on source monitoring tasks; the children substantiated most of their claims to previously assented false items by saying that they really happened. Thus while drawing helps children to remember the reminders, it does not reliably help them to distinguish between what did or did not really happen.

These results have applied as well as theoretical implications for the field of children's suggestibility, interviewing, and memory. On the applied side, they are important only to the degree that they inform us about actual practices of how investigators and therapists use drawings. In this section we will not comment on those practices where drawing is used as a projective technique to identify such conditions as sexual abuse (e.g., Burgess et al., 1981; Riordan & Verdel, 1991) because none of the existing data provide validation for these practices (see Gross & Hayne, 1998); nonetheless, it should be noted that the poor validity may reflect the finding that under certain circumstances children will produce drawings of false events. Rather, we focus on procedures where drawing is used to facilitate verbal communication in interviews with children who are not forthcoming because of cognitive, developmental, or emotional factors (Rae, 1991; Wiehe, 1996). The recommended procedures involve asking the child to draw a picture (of the traumatic event) and, either during or immediately after, asking the child direct questions about the traumatic events about which the interviewer has some *a priori* knowledge.

For example, Malchiodi (1997) recommends drawing as a technique to prompt children's disclosure of sexual abuse and as a way to uncover the details of the abuse. Malchiodi indicates that engaging the child in drawing while asking direct questions about sexual abuse may act as a diversion from the anxiety and discomfort these questions may create, but also facilitate communication, promote self-expression, and absorb anxiety. Pynoos and Eth (1986) include drawing as part of a specialized interview procedure for interviewing child eyewitnesses of traumatic or violent events. The child is asked to draw a picture of something that tells a story. The interviewer then asks general or specific questions about the child's drawing (and story) to link some aspect of the drawing/story directly to the traumatic event. (Note: Before the interview, interviewers are encouraged to learn as much as possible about the traumatic event so that they can be alerted to important omissions in the child's account.) Similar procedures are touted by Burgess and Hartman (1993), who recommend that children be directly asked to draw a picture of "what happened" in the alleged event and then be asked specific questions about the drawing, such as what each person is thinking, what each person is feeling, and what happened to each person.

These procedures raise several concerns that using drawing as an assessment

technique may increase the suggestiveness of the interview and thus children's inaccurate reporting. First, it is never clearly specified what is meant by "specific questions," nor are interviewers warned about the dangers of interjecting new information in conversation or into questions. If interviewers are provided with some information about the alleged event before the interview, and they use this information to question a child, the assessment becomes a suggestive interview, especially if the child is asked by the interviewer to draw details that were omitted.

There is one further concern that is not addressed in the professional literature. Often children draw pictures that are imaginary. It is not known the degree to which children accurately represent reality when they are asked to draw about an "alleged event." That is, it is possible that if they are asked to draw a picture about a time that someone hurt them, they might make up a picture more easily than they would make up a verbal narrative. An anecdote from a clinical interview attended by one of the authors of this paper (M.B.) demonstrates how important it is for interviewers to ask for verbal clarification. A 6-year-old child was evaluated for alleged sexual abuse by family members. During the course of the interview, she was asked to draw her favorite friend. She drew a picture of a dog and with questioning proceeded to provide the dog's name, his age, when she walked him, etc. Knowing the family background, I was surprised that they could take care of a dog and so after the interview I asked the clinician to verify this fact with the mother, who in fact denied the existence of any dog. Of course there is need for experimental support for this anecdote. Nonetheless, it is interesting that none of the articles that we reviewed on the use of art in the assessment process included a procedure wherein the child is asked whether what was depicted in a drawing really happened or whether it was just imaginary.

The results of this study also raise and address several theoretical issues that deserve to be addressed in future research. First, what are the mechanisms that account for superior rehearsal effects and increased misinformation in the drawing compared to the question condition? Although this study was not designed to examine this issue, there are a number of plausible hypotheses. It could be that drawing promotes imagining, which has been related to the creation of false reports in young children (e.g., Ceci, Crotteau-Huffman, Smith, & Loftus, 1994; Ceci, Loftus, et al., 1994), such that the images or thoughts that the child has created during the drawing become confused with the actually experienced event; just talking about events as in the question condition does not promote the same level of imagery. Another hypothesis is that it is motoric involvement that makes the reminders more memorable. It is also possible that it is the visual properties of the drawing that promote later recall (e.g., see Foley et al., 1987); this hypothesis was not supported, however, by the post hoc analysis showing no association between the visual properties of the children's drawings and their recall. Finally, although the drawing and question interviews lasted for the same amount of time, it is clear that children in the drawing condition spent more time



processing (drawing) the reminder. This additional time may have promoted greater recall. Further work to explore the contribution of these various factors will provide a greater understanding of the cognitive bases of suggestibility as well as autobiographical recall in young children.

Next, the false reports provided by the children in the recognition phase seem to be false beliefs (memories). That is, these children came to believe that the false reminders actually happened. Evidence for this assertion is based on the source monitoring data. Recall that children were asked whether their inaccurate assents to false reminded items actually happened. In most cases (86% across conditions), the children claimed that the false reminder *had* actually happened. These results add to the existing evidence that when children later recall misinformation it is often the result of actual memory changes, rather than a result of social-compliance factors (e.g., Ackil & Zaragoza, 1998; Ceci, Loftus, et al., 1994; Poole & Lindsay, 1995). In the present study, the relative rate of false beliefs was similar for the drawing and verbal conditions.

In the literature, the general finding is that false beliefs result when memory for the source of the suggestion is lost, allowing the participant to erroneously recall that the false reminder really happened because he or she failed to remember that it was familiar merely because it had been reminded or suggested. This was the pattern for the children in the question condition, but it was not the pattern for the children in the drawing condition. In the latter group, these children *did* remember being reminded but they nevertheless believed that all reminders also happened. These data are consistent with our initial speculations that reminders that were drawn would be more memorable and would be better recalled. However, they are not consistent with the speculation that drawing would also help children remember whether the event happened. Perhaps because children drew both true and false reminders they eventually could not sort out which ones were actually true or false and as a result relied on the heuristic "Everything I drew is very familiar, so it must have happened." Finally, these data highlight the confusion of the children in the question condition (the parameters of which are similar to those used in other studies of children's source monitoring) about the sources of the reminders and the potentially damaging effects of traditional means of suggestion on children's reports.

## APPENDIX: TRUE AND FALSE ITEMS IN THE MAGICIAN EVENT

True items	False items
Scene 1: ball trick	
Magician wore a black hat	Magician wore black gloves
Child wore a red magic helper cape	Child wore red magic helper boots
Magician put the red ball in her pocket	Magician put the red ball under the table
Magician waved the wand	Magician dropped the wand

True items	False items
Scene 2: water trick	
Child poured water into the cup	Magician drank the water in the cup
Water spilled in the bowl	Water spilled on the floor
Magician stirred the water with a spoon	Magician stirred the water with her finger
Magician put the cup in the corner	Magician put the cup in the garbage
Scene 3: the fall	
Magician fell on her shoelaces	Magician fell on a toy
Magician laughed when she fell	Magician cried when she fell
Magician hurt her hand	Magician hurt her leg
Magician put a Band-Aid on herself	Child put magician's Band-Aid on
Scene 4: clean-up	
Magician put the tricks away in a box	Child put the tricks away in a box
Magician brushed her magic stage with a little brush	Magician brushed her hair with a little brush
Magician showed the child a picture of her puppy	Magician sang the child a song
Magician gave the child a good helper award	Magician gave the child a good helper hug

## 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.
- Ackil, J. K., & Zaragoza, M. S. (1998). The memorial consequences of forced confabulation: Age differences in susceptibility to false memories. *Developmental Psychology*, **34**, 1358–1372.
- Baker-Ward, L., Gordon, B. N., Ornstein, P. A., Larus, D. M., & Clubb, P. A. (1993). Young children's long-term retention of a pediatric examination. *Child Development*, **64**, 1519–1533.
- Brady, M. S., Poole, D. A., Warren, A. R., & Jones, H. R. (1999). Young children's responses to yes–no questions: Patterns and problems. *Applied Developmental Science*, **3**, 47–57.
- Brennan, K. H., & Fisher, R. P. (1998). *Drawing as a technique to facilitate children's recall*. Manuscript in preparation. Results summarized in Poole, D. A., & Lamb, M. E. (1998). *Investigative interviews of children: A guide for helping professionals*. Washington, DC: American Psychological Association.
- Bruck, M., & Ceci, S. J. (1996). *Issues in the scientific validation of interviews with young children: Commentary on Steward, Steward, Farquhar, Myers, & Reinhart "Interviewing young children about body touch and handling."* Monographs of the Society for Research on Child Development (No. 248).
- Bruck, M., Ceci, S. J., & Francoeur, E. (2000). A comparison of three- and four-year-old children's use of anatomically detailed dolls to report genital touching in a medical examination. *Journal of Experimental Psychology: Applied*, **6**, 74–83.
- Bruck, M., Ceci, S. J., Francoeur, E., & Renick, A. (1995). Anatomically detailed dolls do not facilitate preschoolers' reports of a pediatric examination involving genital touching. *Journal of Experimental Psychology: Applied*, **1**, 95–109.
- Burgess, A. W., & Hartman, C. R. (1993). Children's drawings. *Child Abuse & Neglect*, **17**, 161–168.
- Burgess, A. W., McCausland, M. P., & Wolbert, W. A. (1981). Children's drawings as indicators of sexual trauma. *Perspectives in Psychiatric Care*, **19**, 50–58.

- Butler, S., Gross, J., & Hayne, H. (1995). The effect of drawing on memory performance in young children. *Developmental Psychology*, **31**, 597–608.
- Cassidy, D. J., & DeLoache, J. S. (1995). The effect of questioning on young children's memory for an event. *Cognitive Development*, **10**, 109–130.
- Ceci, S. J., & Bruck, M. (1993). Suggestibility of the child witness: A historical review and synthesis. *Psychological Bulletin*, **113**, 403–439.
- Ceci, S. J., & Bruck, M. (1995). *Jeopardy in the courtroom: A scientific analysis of children's testimony*. Washington, DC: American Psychological Association.
- Ceci, S. J., Crotteau-Huffman, M., Smith, E., & Loftus, E. W. (1994). Repeatedly thinking about non-events. *Consciousness and Cognition*, **3**, 388–407.
- Ceci, S. J., Loftus, E. W., Leichtman, M., & Bruck, M. (1994). The role of source misattributions in the creation of false beliefs among preschoolers. *International Journal of Clinical and Experimental Hypnosis*, **62**, 304–320.
- DeLoache, J. S., Anderson, K., & Smith, C. (1995, March). *Interviewing children about real-life events*. Paper presented at the meeting of the Society for Research in Child Development, Indianapolis, IN.
- DeLoache, J. S., & Marzolf, D. P. (1995). The use of dolls to interview young children: Issues in symbolic representation. *Journal of Experimental Child Psychology*, **60**, 155–173.
- Foley, M. A., Aman, C., & Gutch, D. (1987). Discriminating between action and memories: Children's use of kinesthetic cues and visible consequences. *Journal of Experimental Child Psychology*, **44**, 335–347.
- Foley, M. A., Durso, F. T., Wilder, A., & Friedman, R. (1991). Developmental comparisons of explicit versus implicit imagery and reality monitoring. *Journal of Experimental Child Psychology*, **51**, 1–13.
- Gee, S., & Pipe, M.-E. (1995). Helping children to remember: The influence of object cues on children's accounts of a real event. *Developmental Psychology*, **31**, 746–758.
- Goodman, G. S., & Aman, C. (1990). Children's use of anatomically detailed dolls to recount an event. *Child Development*, **61**, 1859–1871.
- Gross, J., & Hayne, H. (1998). Drawing facilitates children's verbal reports of emotionally laden events. *Journal of Experimental Psychology: Applied*, **4**, 163–179.
- Hyman, I. E., Jr., Gilstrap, L. L., Decker, K., & Wilkinson, C. (1998). Manipulating remember and know judgements of autobiographical memories: An investigation of false memory creation. *Applied Cognitive Psychology*, **12**, 371–386.
- Hyman, I. E., Jr., & Pentland, J. (1996). The role of mental imagery in the creation of false childhood memories. *Journal of Memory and Language*, **35**, 101–117.
- Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. *Psychological Bulletin*, **114**, 3–28.
- Lamb, M. E., Hershkowitz, I., Sternberg, K. J., Boat, B., & Everson, M. D. (1996). Investigative interviews of alleged sexual abuse victims with and without anatomical dolls. *Child Abuse & Neglect*, **20**, 1251–1259.
- Malchiodi, C. A. (1997). *Breaking the silence: Art therapy with children from violent homes* (2nd ed.). Bristol, PA: Brunner/Mazel.
- McAuliff, B., Kovera, M., & Viswesvaran, C. (1998, March). *Methodological issues in child suggestibility research: A meta-analysis*. Paper presented at the meeting of the American Psychology–Law Society, Redondo Beach, CA.
- Oppawsky, J. (1991). Utilizing children's drawings in working with children following divorce. *Journal of Divorce & Remarriage*, **15**, 125–141.
- Paivio, A. (1968). A factor-analytic study of word attributes and verbal learning. *Journal of Verbal Learning and Verbal Behavior*, **7**, 41–49.
- Parker, J. F. (1995). Age differences in source monitoring of performed and imagined actions on immediate and delayed tasks. *Journal of Experimental Child Psychology*, **60**, 84–101.

- Peterson, C., & Bell, M. (1996). Children's memory for traumatic injury. *Child Development*, **67**, 3045–3070.
- Peterson, C., & Biggs, M. (1997). Interviewing children about trauma: Problems with "specific" questions. *Journal of Traumatic Stress*, **10**, 279–290.
- Poole, D. A., & Lamb, M. E. (1998). *Investigative interviews of children: A guide for helping professionals*. Washington, DC: American Psychological Association.
- 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.
- Poole, D. A., & Lindsay, D. S. (in press). Effects of parents' suggestions, interviewing techniques and age on young children's reports. *Journal of Experimental Psychology: Applied*.
- Poole, D. A., & White, L. T. (1991). Effects of question repetition on the eyewitness testimony of children and adults. *Developmental Psychology*, **27**, 975–986.
- Pynoos, R. S., & Eth, S. (1986). Witness to violence: The child interview. *Journal of the American Academy of Child Psychiatry*, **25**, 306–319.
- Rae, W. A. (1991). Analyzing drawings of children who are physically ill and hospitalized, using the ipsative method. *Children's Health Care*, **20**, 198–207.
- Riordan, R. J., & Verdel, A. C. (1991). Evidence of sexual abuse in children's art products. *The School Counselor*, **39**, 116–121.
- Salmon, K., Bidrose, S., & Pipe, M.-E. (1995). Providing props to facilitate children's event reports: A comparison of toys and real items. *Journal of Experimental Child Psychology*, **60**, 174–194.
- Salmon, K., & Pipe, M.-E. (1997). Props and children's event reports: The impact of a one year delay. *Journal of Experimental Child Psychology*, **65**, 261–292.
- Salmon, K., & Pipe, M.-E. (2000). Recalling an event one year later: The impact of props, drawing and a prior interview. *Applied Cognitive Psychology*, **14**, 99–120.
- Saywitz, K., Goodman, G., Nicholas, G., & Moan, S. (1991). Children's memory for genital exam: Implications for child sexual abuse. *Journal of Consulting and Clinical Psychology*, **59**, 682–691.
- Steward, M. S., & Steward, D. S. (1996). Interviewing young children about body touch and handling. *Monographs of the Society for Research in Child Development*, **61** (4–5, Serial No. 248).
- Walker, N., Lunning, S., & Eilts, J. (1996, June). *Do children respond accurately to force choice questions?* Paper presented to the NATO Advanced Study Institute: Recollections of Trauma: Scientific Research and Clinical Practice, Talmont Saint Hilaire, France.
- Warren, A., Woodall, C. E., Hunt, J. S., & Perry, N. W. (1996). "It sounds good in theory, but . . .": Do investigative interviewers follow guidelines based on memory research? *Child Maltreatment*, **1**, 231–245.
- White, S. (1990). The investigatory interview with suspected victims of child abuse. In A. M. LaGreca (Ed.), *Through the eyes of the child: Obtaining self-reports from children and adolescents* (pp. 368–394). Needham Heights, MA: Allyn & Bacon.
- Wiehe, V. (1996). *Working with child abuse and neglect*. Thousand Oaks, CA: Sage.
- Wilson, J. C., & Pipe, M.-E. (1989). The effects of cues on young children's recall of real events. *New Zealand Journal of Psychology*, **18**, 65–70.

Received August 25, 1998; revised December 20, 1999