“I saw it with my own ears”: The effects of peer conversations on preschoolers’ reports of nonexperienced events

Gabrielle F. Principe * and Stephen J. Ceci

Department of Human Development, Cornell University

Received 24 April 2002; revised 21 August 2002

Abstract

The study was designed to explore the effects of naturally occurring peer interactions and repeated suggestive interviews on preschoolers’ (N = 96, Mean age = 54 months) memories for a personally experienced event, namely a staged archaeological dig. During the dig, one third of the children witnessed two “target” activities. A second third of the children were the classmates of those in the first group, but did not witness the target activities. The remaining children were not the classmates of those who witnessed the target activities, nor did they witness the target activities themselves, and thus served to provide a baseline against which to assess the effects of peer contact. Following the dig, the children were interviewed in either a neutral or suggestive manner on three occasions. Results from a fourth interview by a new examiner revealed that the combination of suggestive interviews and peer exposure led to claims of witnessing the target activities by the classmate group that were comparable to the children who actually did witness these activities. Further, assent rates to misleading questions employing peer pressure and false claims of actually seeing versus merely hearing about the target activities were elevated following opportunities to discuss these activities with peers.

© 2002 Elsevier Science (USA). All rights reserved.

Keywords: Memory; Peers; Suggestibility; Legal testimony; Interviews; Childhood
In response to the increased frequency with which young children are participating as witnesses in the legal system (see Poole & Lamb, 1998), much attention has been focused recently on factors affecting the reliability of children’s reports of events. This line of work has demonstrated that children can be either greatly inaccurate or highly accurate in relating their experiences. One the one hand, research on suggestibility has found that coercive interviewing procedures can produce serious degradations in children’s accuracy. For example, when exposed to repeated misleading interviews incorporating stereotypes (Leichtman & Ceci, 1995) or guided imagery (Ceci, Crotteau-Huffman, Smith, & Loftus, 1994a), children can be persuaded to report false events with such a degree of credibility that even trained professionals cannot tell whether the events in question did or did not take place. On the other hand, the event memory literature has shown that even very young children can display impressive autobiographical memories when they are competently interviewed. For instance, Ornstein and his colleagues (e.g., Baker-Ward, Gordon, Ornstein, Larus, & Clubb, 1993; Ornstein, Gordon, & Larus, 1992; see Ornstein, Baker-Ward, Gordon, & Merritt, 1997, for a review) have shown that when questioned in a neutral and supportive manner, children as young as 3 and 4 years of age can provide accurate and surprisingly rich accounts of personal experiences.

Converging findings from these two research traditions makes it apparent that children’s reporting accuracy is largely dependent on the interviewer’s interrogation style, such that one has greater confidence in the accuracy of statements elicited in a neutral context than those made only after multiple suggestive interviews. Indeed, some experts have argued that suggestive questioning practices are a greater source of false responses than any limitations inherent in young children’s cognitive abilities (Ceci, Bruck, & Battin, 2000; Lamb, Sternberg, Esplin, & Orbach, 1997; McGough, 1996). Thus, the problem has shifted from children’s causality for erroneous reports to interviewers’ causality. Moreover, the detrimental effects of suggestive interviewers are thought to be particularly acute for preschool age children as the literature on context specificity suggests that their memory performance is especially dependent on aspects of the context in which remembering is assessed (Folds, Footo, Guttentag, & Ornstein, 1990; Ornstein, Baker-Ward, & Naus, 1988). These characteristics of young children’s reporting abilities have prompted a number of research groups and professional organizations to construct interviewing guidelines for those who question child witnesses—guidelines that, in part, are designed to increase the accuracy of memory reports by eliminating suggestive techniques (see e.g., American Professional Society on the Abuse of Children, 1990; Gordon, Schroeder, Ornstein, & Baker-Ward, 1995; Home Office, 1992; Poole & Lamb, 1998; Sorensen, Bottoms, & Perona, 1997; Yuille, 1988).

Admittedly, however, the use of nonsuggestive questioning procedures may not eliminate all concerns about children’s testimonial accuracy. This
is because young witnesses are exposed to a range of experiences outside of the formal interview environment that have the potential to interfere with subsequent recall. Indeed, in many cases, doubts about a child’s reliability may not stem from earlier interviews per se, but rather from material gained from other sources. For example, recent research has shown that postevent exposure to misleading information from parents (Poole & Lindsay, 1995, 2001, 2002) and on television (Principe, Ornstein, Baker-Ward, & Gordon, 2000) can induce false reports of nonexperienced events, even when interviews are highly supportive of correct recall. Importantly, in the investigations by Poole and Lindsay (1995, 2001, 2002) and Principe et al. (2000), the children were never told that the nonexperienced events actually happened to them, nor were they encouraged by their parents or the experimenters to report the suggestions; rather the children simply were exposed to information prior to the memory assessment that was inconsistent with the details of an already experienced event. Thus, these studies demonstrate that incidental experiences can compromise the accuracy of children’s accounts, and as such, highlight the importance of examining the mnemonic consequences of naturally occurring events that have the potential to contaminate already stored memories.

One source of potential misinformation that has not been studied systematically is peer conversations. A compelling reason for examining the influence of peer conversations on remembering concerns the social nature of autobiographical memory (Neisser, 1988; Nelson, 1993). Children share their memories with others through conversation. In the midst of this exchange of memories, children are constantly encountering others’ versions of the past. When there is disagreement or uncertainty about what happened, children may be prone to making “source” errors if they confuse the mental images inspired by another child’s account with the images generated during an actual experience. In some cases, children may immediately adopt another child’s construal of an event. In other cases, children may retain both interpretations for a while, but later misattribute the other child’s recollection as their own.

A number of researchers have argued that source monitoring errors are a primary cause of suggestibility effects—children misidentify the memory records of postevent suggestions as memories of the experienced event itself (Leichtman, Morse, Dixon, & Spiegel, 2000; Lindsay & Johnson, 1987). However, much of this work has focused on the role of suggestions during interactions with adults, such as exposure to questioning that inspires the visualization of nonexperienced events, conversations that lead to the formation of strong stereotypes, and storybook reading that describes both occurring and nonoccurring activities (Ceci et al., 1994a; Ceci, Loftus, Leichtman, & Bruck, 1994b; Leichtman & Ceci, 1995; Poole & Lindsay, 2001, 2002). Relatively little is known about the extent to which the suggestions of peers can create source failures.
Examining peer conversations is an important issue not only for the study of memory, but also for applications in the research literature in efforts to understand the testimony of young children. This is because child witnesses presumably talk with their friends about the incidents in question. Over time, such discussions may cause children within a peer group to hold increasingly similar hypotheses about what happened. An historical demonstration of this phenomenon was provided by Sjöberg (1995, 1997) in his description and analysis of the Great Swedish Witch Panic of the 17th century. Between 1670 and 1671, testimonies were taken from 588 children concerning their alleged contacts with witches in the Swedish village of Rättvik. The names and ages of the children were recorded in parch records, along with the circumstances under which the children’s testimonies were given—in particular, whether they were given in private or during group prayer meetings after the children had been standing in line with their peers. Sjöberg showed that it was statistically more likely for children to claim evidence of satanic practice and abduction if their testimony was taken after they had stood in line with peers than if they had been interviewed in the absence of peer experience. The suggestion is that standing in line exposed the children to witchcraft stories by other children, and hence led to broader allegations that contained similar themes.

Concerns about peer influence are especially relevant for cases involving multiple purported victims because fact finders may rely on the number and similarity of allegations to determine the credibility of any single child’s testimony. Indeed, it is likely very compelling when judges and juries hear child after child tell the same story, especially if they believe that each witness has arrived at the same storyline independently. But at least in some situations, the exact opposite might be the case. The story may have been arrived at in a collaborative manner among peers who initially had very different memories about what took place. As such, it is important to examine the extent to which children’s reports of shared events become increasingly similar over time.

A related concern is the effects of peer pressure on children’s testimonial accuracy. Garven, Wood, and Malpass (2000) found that telling children what their friends had replied failed to persuade them to provide false accounts of events consistent with the alleged experiences of peers. However, Garven and colleagues used a mild form of peer pressure during an otherwise nonsuggestive interview, indicating that peer pressure, in and of itself, may not be powerful enough to promote reporting of nonexperienced events. To test this hypothesis, Bruck, Ceci, and Hembrooke (in press)

---

1 Albeit anecdotal, numerous reports of jurors have expressed the view that they believed the children’s allegations were true precisely because so many children testified with similar details, leading jurors to convict because of the implausibility of so many children making up the same false story (see Ceci & Bruck, 1995, for examples from multivictim–multiperpetrator cases).
examined the effects of peer pressure when included in repeated, aggressively coercive interviews. Bruck et al. observed that a substantial number of children assented to having experienced a previously denied event following questions embedded in highly suggestive interviews that insinuated that their friends had already told.

Although children testify in a range of criminal and civil cases, research on children’s memory and suggestibility has its largest impact in cases involving allegations of sexual and physical abuse, especially cases in day-care settings (Ceci & Bruck, 1995). Peer influence is a particularly serious problem for day-care abuse investigations because such cases usually involve multiple alleged victims who spend time together on a daily basis. Further, case studies reveal that child witnesses in day-care cases often undergo group counseling sessions with their classmates, thus providing children additional opportunities to hear and be influenced by the perspectives of their peers (Ceci & Bruck, 1995). In other cases, child witnesses are encouraged to disclose the details of their presumed victimization by legal and mental health professionals who tell them what their peers allegedly have reported (Ceci & Bruck, 1995).

To illustrate that discussions among peers can influence the testimony of children, consider the following exchange between a child witness and a forensic investigator in the highly publicized Wee Care Nursery School case in Maplewood, New Jersey. In this case, Kelly Michaels, a teacher at the school, was accused of sexually abusing her students. Here a child reveals that the source of her knowledge was another child rather than her own observation:

*Interviewer: Do you know what [Kelly] did?*

*Child: She wasn’t supposed to touch somebody’s body. If you want to touch somebody, touch your own.*

*Interviewer: How do you know about her touching private parts? Is that something you saw or heard?*

*Child: Max told me.*

(Ceci & Bruck, 1995, p. 150)

The following excerpt, taken from the trial testimony of the Little Rascals Day Care case in Edenton, North Carolina, provides an additional example of a child arriving at a conclusion about what actually occurred on the basis of statements made by peers:

*Child: He touched kids. I just know it because if he’s touched everybody else and everyone else has been up here then I kind of know that they’ve been touched.*

*Attorney: Okay. Did somebody tell you that?*

*Child: No sir. It’s just that everybody has been up there and I think they have been telling the truth.*

(Ceci & Bruck, 1995, pp. 150–151)
A final example from the Wee Care Nursery School case provides a potent demonstration that investigators in actual cases, at least at times, have tried to persuade children to disclose by telling them what their peers allegedly had reported:

*Interviewer:* All the other friends I talked to told me everything that happened. Randy told me. Charlie told me, Connie told me... And now it’s your turn to tell. You don’t want to be left out, do you? (Ceci & Bruck, 1995, p. 149)

The first scientific study of peer suggestibility was conducted by Small (1896). Small uncorked a bottle of clear liquid and asked children to raise their hand when they could smell the fragrance emitted from the bottle. The bottle contained water, yet many of the children claimed to smell its fragrance. Further, Small found that many of the children were influenced by their peers’ responses—upon seeing one peer raise his or her hand, others raised their hands, too.

Although Small’s study was seminal, the most widely cited study of peer suggestibility is Binet’s (1900) investigation. After showing children a series of objects, Binet asked several misleading questions about the objects and told the children to call out the answer to each question as quickly as possible. Binet found that the children who replied second and third often gave the same answer as the first respondent, even if the answer was erroneous.

For nearly 100 years, the work of Small (1896) and Binet (1900) were the sole studies of the effects of peer influence on children’s suggestibility. Recently, however, some anecdotal accounts have provided evidence that children will fabricate reports of events that they did not witness but may have found out about from peers. Pynoos and Nader (1989) examined children’s memory for a sniper attack at their school during which one child and a bystander were killed. When later questioned, some of the children who were not at school during the shooting, including those already at home or on vacation, produced quite elaborate narratives describing the attack. Presumably, these children heard about the attack from their peers who were present and later constructed their accounts on the basis of this secondhand information.

In an unpublished study by Pettit, Fegan, and Howie (1990), two unfamiliar adults accidentally knocked over a cake during a visit to a preschool classroom. When later questioned, six of the seven children who were absent from school during the adults’ visit indicated that they had been present. These children may have provided erroneous reports so that they would feel as if they were part of the same group as their friends who actually witnessed the event.

Although these more recent accounts demonstrate that the peer groups’ experiences can lead non-participants to produce false reports, the effect of peer interaction on non-participants’ accounts remains unclear. The next step is to experimentally manipulate the participation and non-participation
of children within peer groups in the to-be-remembered event. Further, prior work does not address the extent to which exposure to intervening suggestive interviews might augment the influence of peers. Indeed, this question has special relevance to cases involving allegations of child sexual abuse as analyses of professional field interviews have shown that police and social service officials at least at times have used highly suggestive techniques in attempts to elicit disclosures of sexual abuse from young witnesses (see e.g., Lamb, Hershkowitz, Sternberg, Boat, & Everson, 1996a; Lamb et al., 1996b; Sternberg et al., 1996; Warren et al., 1999). The present study offers an analysis of these issues by examining the impact of naturally occurring interactions with peers and exposure to repeated suggestive interviews on children’s reporting accuracy.

The Present study

The goal of this investigation was to build on suggestibility research to examine the combined effects of peer influence and biased questioning on children’s memory for another person’s behaviors. Of particular interest was the extent to which the peer group’s actual experiences in an event might lead non-participants to produce false reports of another person’s behavior, especially when children are exposed to suggestive interviews that are inconsistent with their own experiences but consistent with the peer group’s experiences.

To examine these issues, three groups of preschool-aged children participated in a staged classroom event. This event was designed to be interactive, enjoyable, and hence highly memorable, so as to maximize the likelihood that the children would later jointly reminisce about the details of the activity. During the event, one third of the children, those in the Witness condition, witnessed two extra activities, denoted as the “target” activities. A second third of the children, those in the Classmate condition, were the classmates of those in the first group, but did not witness the target activities. It was expected that some of these children would receive indirect information about the target activities through naturally occurring conversations with some of their classmates who actually did witness these activities. The remaining children in the Control condition were not the classmates of those who witnessed the target activities, nor did they witness the target activities themselves; they were included to provide a baseline against which the effects of peer contact could be assessed by providing the likelihood that a target activity would be reported by a random non-participant without any exposure to peer participants.2

2 It was unnecessary to include a fourth experimental group, one that witnessed the target activities but who were not exposed to classmates who did not, because such a condition would be unrevealing in terms of the key questions of the study.
The children in all three experimental conditions were questioned about the details of the event on three different occasions spread out over a 3-week time period. Half of the children in each of the three groups were questioned in a suggestive manner, whereas the remaining half received neutral interviews. Embedded in the suggestive interviews were strongly worded leading questions that suggested that the interviewee had witnessed the two target activities, when in fact only one third of these children actually saw these activities take place. Note, however, that these questions did not implicate that peers also had witnessed the two target activities. For the children who did witness the target activities, these questions were expected to help reinforce their recollections of these activities, thus increasing the likelihood that they would retain their memories for the activities, and also talk with some of their classmates about their occurrences.

To assess the consequences of peer interaction and suggestive questioning, all of the children received a fourth interview by a new examiner four weeks after the initial classroom event. Included in this final assessment were source monitoring questions that were designed to probe the origin, or the source, of reports of target activities. Of particular interest was the extent to which peer conversation would elevate source confusions. Also embedded in the final interview were peer conformity, or peer pressure, probes. These questions were included to examine the effects of peer conversations on letting children know that their friends have “already told.”

Method

Participants

Ninety-six 3.5- to 5-year-old children (Mean age = 54 months, Range = 42–60 months) participated in the experiment. Approximately half of the children were female (55%). The children were recruited from preschools and day-care centers in the Finger Lakes region of New York State. Reflecting the population of these child-care centers, 81% of the children were European American, 9% were African American, and 10% were Asian American.

Experimental design

The participants were assigned to one of six conditions: one of three conditions that differed in the degree of participation in the target activities (Witness versus Classmate versus Control), and within each of these there were one of two conditions that differed in terms of the type of intervening interviews (Suggestive versus Neutral). The groups ranged in size from 16 to 18 and the number of females in each group ranged from 8 to 10. Likewise,
the proportions of children in each racial category were approximately equal across the six experimental groups. Each classroom ($N = 10$) was randomly assigned to participation condition and the children within each classroom were randomly assigned to interview condition.

Procedure

The children were seen during five sessions, each of which took place in a separate room at their child-care centers. The sessions were separated by an interval of approximately 1 week ($Mean = 7$ days, $Range = 4–10$ days) and were videotaped for subsequent coding.

The to-be-remembered event. During the first session, the children, in groups ranging in size from 6 to 8, took part in a contrived dig with a fictitious archaeologist named Dr. Diggs. The role of Dr. Diggs was played by a male or female graduate student. Dr. Diggs led the children through an event in which they used plastic hammers to dig pretend artifacts, such as dinosaur bones, gold coins, and jewels, out of specially constructed blocks made of mortar mix and play sand. Once all of the artifacts were uncovered, Dr. Diggs labeled and passed around each artifact. During the course of his presentation, Dr. Diggs pointed out that two of the artifacts—a map to a buried treasure and a rock with a secret message—were very special. He [She] warned the children that they must be careful not to spill anything on the map because the ink might smear and then he [she] would never be able to find the buried treasure. He [She] also warned the children not to drop the rock because it might break and then he [she] would never be able to find out what the secret message says.

During the course of the archaeological event, the children in the Witness conditions saw Dr. Diggs ruin the two special artifacts (heretofore referred to as the target activities). In particular, Dr. Diggs “accidentally” spilled a drink on the treasure map, smearing the ink and rendering the map illegible. When he [she] did this, Dr. Diggs appeared upset, causing an abrupt halt to the event, and said, “I’ve messed up the map and now I’ll never find the buried treasure!” Also, Dr. Diggs “accidentally” dropped the rock with the secret message and said, “I’ve broken the rock and now I’ll never know what the secret message says!”

Intervening interviews. Starting 1 week following the archeological event, the children were questioned individually by a condition-blind interviewer in either a neutral or suggestive manner on three different occasions spread out over a 3-week period. Both the neutral and suggestive interviews followed a structured format and began with an open-ended question about Dr. Diggs’ visit: “Remember the day that Dr. Diggs visited your classroom? Well, I wasn’t there that day, and I’d like you to tell me everything that happened.” This was followed by additional open-ended prompts (e.g., “What else happened?”) until the child’s recall was exhausted.
In addition, for the children in the neutral interview groups, target events not mentioned in response to the open-ended probes were queried with specific probes: “I heard something about a treasure map [rock with a secret message]. Do you know anything about that?” For each target activity reported at the open-ended or specific level of questioning, the children were asked to elaborate (e.g., “Tell me more about that.”). One half of these children were questioned first about the treasure map and then about the secret message rock, whereas the other half were probed in the reverse order.

In contrast, once open-ended recall was exhausted, the children in the suggestive interview conditions were provided with two sets of suggestions, one having to do with Dr. Diggs spilling a drink on the treasure map and the other having to do with Dr. Diggs breaking the rock with the secret message (e.g., “What did Dr. Diggs spill on the treasure map, did he [she] spill a cup of coffee, or a cup of juice?” and “How did Dr. Diggs break the rock, did he [she] step on it, or did he [she] drop it?”). The exact suggestive questions about the target activities were different during each of the interviews, but each had similar implications (i.e., spilling a drink on the treasure map and breaking the rock with the secret message). Further, the children were asked to elaborate after each suggestive question (e.g., “Tell me more about that.”). Note that these suggestions were consistent with what the children in the Witness conditions experienced, but were misleading for those in the remaining groups. Four parallel protocols were used for these children that differed only in the order of the suggestive questions, and one quarter of the participants received each version of the interview.

Final interview. Approximately four weeks after the dig, a new condition-blind interviewer questioned the children. Four comparable protocols were used that differed only in the order of questioning, and one quarter of the participants received each version of the interview.

This final assessment followed a structured format that involved five levels of questions: (a) open-ended questions, (b) specific questions, (c) source questions, (d) countersuggestion questions, and (e) peer conformity questions. After establishing rapport and orienting the child to the archeological event, the interview began with an open-ended question about Dr. Diggs’ visit: “Remember the day that Dr. Diggs visited your classroom? Well, I wasn’t there that day, and I’d like you to tell me everything that happened.” Once open-ended recall was exhausted, target activities not already reported were queried with specific probes: “Did anything happen to a treasure map?” and “Did anything happen to a rock with a secret message?” For each target activity mentioned in response to an open-ended or specific probe, the children were asked to elaborate (e.g., “Tell me more about that.”).

For each target activity reported at the open-ended or specific level of questioning, the children were given an additional question to probe the
source of their memory for that activity. These questions, referred to as the source questions, asked whether the children merely “heard” about the target activity, or actually saw Dr. Diggs do the target activity: “Did you see Dr. Diggs spill his [her] drink on the treasure map [break the rock with the secret message] with your own eyes, or hear that he [she] did it?”

Finally, following the oppositional logic of Leichtman and Ceci (1995), the subset of children who reported that they saw Dr. Diggs do a target activity were asked a countersuggestion that insinuated that the interviewee had not witnessed the target activity: “You really didn’t see him [her] spill his [her] drink on the treasure map [break the rock with the secret message], did you?” The purpose of the countersuggestion questions was to gauge the strength of the children’s belief that they actually witnessed the activity in question.

Further, to examine effects of letting children know that their friends had “already told,” the participants who had not already reported one or both of target activities were asked a question that insinuated that some of their classmates already had reported a ruined treasure map or a broken rock (heretofore referred to as the peer conformity questions): “Some of your friends told me that Dr. Diggs spilled his [her] drink on the treasure map [broke the rock with the secret message]. Now it’s your turn to tell. Did he [she] do that?”

Coding of the interviews

The videotapes of the final interviews were transcribed and coded by condition-blind raters. The data were coded for both the target activities reported (i.e., Dr. Diggs spilling his [her] drink on the treasure map and Dr. Diggs breaking the secret message rock) and the type of probe needed to elicit the information (i.e., open-ended versus specific versus peer conformity). In addition, when asked, the children’s responses to the witness and countersuggestion questions were scored.

Following Bruck et al. (in press), the children’s narrative responses to open-ended and specific probes about the target activities were characterized in terms of length or the number of “utterance ideas.” An utterance idea was considered to be any meaningful statement, usually but not always containing a verb and bounded by a pause, that had not been mentioned previously by either the interviewer or the child. Examples of utterance ideas included, “The rock broke,” “He messed everything up,” “She did it in the lounge,” and “Ink all over it.” (in response to the specific question, “Did anything happen to a treasure map?”). Utterances that did not concern the target activities or for which their accuracy could not be determined were not coded.

Utterance ideas regarding the target activities also were classified as either “Witnessed-Suggested” or “Constructive.” Witnessed-Suggested responses
referred to details that were part of the standard script for the children in the Witness groups or those that were reproduced literally as they had been described in the suggestive interviews. For example, if a child said, “Dr. Diggs will never find the treasure now,” this information would be scored as Witnessed-Suggested because Dr. Diggs actually stated in the Witness conditions that the ruined map would prevent him from finding the buried treasure. Likewise, the statement, “Dr. Diggs spilled coffee on the map,” would be counted as Witnessed-Suggested because these details were insinuated in one of the questions asked during the suggestive interviews. In contrast, statements coded as Constructive involved going beyond the information experienced by the children in the Witness conditions or provided in the suggestive interviews. Thus, if a child stated, “Dr. Diggs dried off the map,” these nonoccurring and nonsuggested details would be coded as Constructive because they reflect an embellishment of the suggested information that Dr. Diggs spilled a drink on the map. Because there was a good deal of overlap between what the children in the Witness conditions experienced and what was contained in the suggestive interviews, utterance ideas could not be separated in terms of their source (i.e., the actual experiences of the children in the Witness conditions versus the material contained in the suggestive interviews).

Of the 268 interviews coded, approximately 20% (n = 54) were coded independently by two judges and checked for inter-rater reliability. Inter-rater agreement as measured by $\kappa$ was quite good, ranging from .95 to 1 for the target activities reported, the type of probe needed to elicit the information, the persistence of the report following source and countersuggestion questions, and the number of utterance ideas. Coding discrepancies were mainly oversights that were resolved through discussion.

Teacher questionnaires

As a manipulation check, during the duration of the study, the teachers in Witness and Classmate classrooms were asked to note overheard, naturally occurring discussions between their students regarding the target activities. However, the teachers were discouraged from initiating or participating in conversations about the archeological event. Using a standardized form, the teachers recorded the content and date of the conversations. The mean number of conversations about target activities recorded within each classroom was 5 (Range = 2–11), and the majority of conversations (89%) took place during the first week following the archaeological event. Despite their relatively low frequency, these conversations between classmates suggest that numerous additional conversations about the target activities most likely occurred among classmates but were not overheard or recorded by their teachers. Thus, these data document that some relevant, naturally occurring conversations among peers did take place.
Results

The general analytic strategy involved a series of 2 (intervening interview condition: Neutral versus Suggestive) × 3 (participation condition: Witness versus Classmate versus Control) analyses of variance (ANOVAs) to examine the effects of the experimental manipulations on the children’s reports of the target activities (i.e., spilling a drink on the treasure map and breaking the secret message rock) during the final interview. In these analyses, the type of intervening interview and the level of participation were between-subjects factors. Only significant main effects and interactions are reported.

After a brief treatment of preliminary analyses, the proportion of target activities reported in response to open-ended and specific questions is summarized. Examined next are the children’s responses to source questions and countersuggestions following reports of target activities. Additional analyses focus on the degree to which the subset of children who failed to report a target activity at the open-ended or specific levels of questioning subsequently assented to this activity following peer conformity questions. Finally, differences in the length and content of true and false narratives of the target activities are explored.

Preliminary analyses

Preliminary analyses indicated no main effects or interactions in memory performance as a function of the particular confederate who played the role of “Dr. Diggs,” the order of questioning during the intervening and final interviews, the individuals who served as the intervening and final interviewers, the various child-care centers, the specific classrooms within the centers, age in months, or gender. Accordingly, the data were collapsed over these variables for subsequent analyses.

Proportion of target events reported

To examine the effects of the experimental manipulations on the children’s recall of the target activities, the proportions of these activities that were reported at the open-ended and specific levels of questioning were calculated and averaged across the children in each of the six experimental groups. Note that because the children in the Witness conditions actually experienced the target activities, their reports of these activities reflected accurate recall. In contrast, because the target activities did not occur during the archaeology dig for the children in the Classmate and the Control conditions, their reports of target activities were inaccurate.

Total recall. The proportion of target activities reported at the final interview is shown in Table 1, as a function of experimental group and level of
recall: open-ended, specific, and total (open-ended plus specific). As can be inferred from the data in the third column, the children’s total recall of the target activities differed as a function of interview condition, with the children in the Suggestive Interview groups ($M = 80$) showing higher levels of recall than those in the Neutral Interview groups ($M = 39$), $F(1, 90) = 37.92$, $p < .0001$. There was also a main effect of participation condition, $F(2, 90) = 23.96$, $p < .0001$, the result of the children in the Witness conditions ($M = 86$) recalling a higher proportion of target activities than those in the Classmate conditions ($M = 62$), who in turn reported more target activities than those in the Control conditions ($M = 29$), $Fs(1, 90) \geq 10.59$, $ps \leq .01$. The interview condition $\times$ participation condition interaction was significant, $F(2, 90) = 5.60$, $p < .01$, indicating that the effects of type of intervening interview differed as a function of the children’s level of participation.

To explore further this interaction, separate follow-up analyses were carried out for each interview condition. Analysis of participation group differences for the neutral interviews resulted in a significant main effect of level of participation, $F(2, 43) = 22.41$, $p < .0001$, with the children in the Witness/Neutral Interview condition recalling a higher proportion of target activities than those in the Classmate/Neutral Interview condition, who in turn reported more target activities than those in the Control/Neutral Interview condition, $Fs(1, 43) \geq 6.43$, $ps \leq .01$.

Follow-up tests among the children in the suggestive interview conditions indicated that the effect of level of participation was significant, $F(2, 47) = 5.70$, $p < .01$, due to the children in the Witness/Suggestive Interview group and those in the Classmate/Suggestive Interview group reporting greater proportions of target activities compared to those in the Control/Suggestive Interview group, $Fs(1, 47) \geq 8.27$, $ps \leq .01$. Further, the recall scores of the children in the Witness/Suggestive Interview condition were not significantly different from those of the children in the Classmate/Suggestive Interview condition, demonstrating that, when exposed to suggestive interviews, the children in the Witness and Classmate conditions were equally likely to report target activities.
The group differences in recall of target activities among the Classmate and Control children demonstrate the powerful effect of peer contact on children’s reports. Because these differences occurred under both interviewing conditions, suggestive interviewing by itself cannot account for the rise in erroneous reports of target activities among the Classmate children relative to the Control children. Rather, peer interaction was the major source of false reports among non-witnesses. Moreover, when the non-witness children were exposed to the combined effects of both classmates who actually did witness the activities, and to suggestive interviews, they ended up reporting as many target activities as the children in the Witness conditions who actually did experience these activities. Thus, it is not merely that peers and suggestive interviewers have deleterious effects on the report accuracy of non-witness children, but the combined effects of peer influence and suggestive questioning actually resulted in levels of recall that were indistinguishable from witness children in terms of magnitude.

Open-ended recall. Because spontaneity is often used in clinical and legal settings as a marker for judging the credibility of children’s accounts (Ceci & Bruck, 1995), it was of interest to examine the extent to which the six experimental groups differed in their reports of the target activities in response to open-ended questions. As can be seen in the data presented in the first column of Table 1, there was a main effect of interview condition, $F(1, 90) = 18.15, p < .0001$, with the children in the Suggestive Interview conditions ($M = .48$) spontaneously reporting a higher proportion of target activities than those in the Neutral Interview conditions ($M = .17$). There was also a significant main effect of participation condition, $F(2, 90) = 5.86, p < .01$. Follow-up contrasts indicated that the children in the Witness conditions ($M = .52$) spontaneously recalled a greater proportion of target activities than those in the Classmate conditions ($M = .34$), who in turn reported a greater proportion of target activities than those in the Control conditions ($M = .12$), $F$s $\geq 3.80, ps \leq .05$. Thus, these data show that the deleterious effects of peer conversations are not limited to cued recall, but also can result in open-ended intrusions of nonexperienced events.

Memory for source of reports of target events

The extent to which the children reported actually witnessing (as opposed to hearing about) the target activities following source questions (e.g., “Did you see Dr. Diggs spill his drink on the treasure map with your own eyes, or hear that he did it?”) and countersuggestions (e.g., “You really didn’t see him spill his drink on the treasure map, did you?”) is shown in Table 2. The first column in Table 2 indicates the proportions of activities reported at the open-ended or specific levels of questioning for which the children recalled as actually witnessed in response to source questions; and the second column represents the proportions of witnessed activities that continued to
be reported as witnessed following countersuggestion questions. As can be inferred from these data, the children’s responses to the source questions varied as a function of the type of intervening interview, \( F(1, 90) = 12.57, p < .001 \), with the children in the Suggestive Interview conditions (\( M = .59 \)) more likely to report witnessing the target activities compared to those in the Neutral Interview conditions (\( M = .34 \)). The children’s responses to source questions also varied in terms of the level of participation, \( F(2, 90) = 30.94, p < .0001 \). Contrasts indicated that the children in the Witness groups (\( M = .81 \)) were much more likely to recall actually witnessing a target activity than those in the Classmate groups (\( M = .41 \)), who in turn were more likely to report witnessing a target activity than those in the Control groups (\( M = .14 \)), \( F(1, 90) \geq 9.03, ps \leq .01 \). Thus, these data demonstrate that interactions with peers, in and of themselves, can elevate claims of actually seeing (as opposed to merely hearing about) non-witnessed events.

Inspection of the third column of Table 2 indicates that the children’s responses to the countersuggestion questions varied by interview condition, \( F(1, 90) = 5.35, p < .05 \), with those in the Suggestive Interview conditions (\( M = .37 \)) more likely to continue to report witnessing target activities following the countersuggestions than those in the Neutral Interview conditions (\( M = .20 \)). There was also a main effect of participation level, \( F(2, 90) = 9.70, p < .001 \). Contrasts indicated that the children in the Witness groups (\( M = .52 \)) were far more likely to persist in reports of a target activity when challenged with a countersuggestion than those in the Classmate (\( M = .21 \)) and Control groups (\( M = .12 \)), who did not differ from one another, \( F(1, 90) \geq 11.35, ps \leq .001 \). This last contrast suggests that for the countersuggestions, the driving source of children’s reports of witnessing versus hearing about the target activities may have been mainly suggestive questioning by interviewers rather than exposure to peer witnesses.

However, in support of the hypothesis that naturally occurring discussions among children can influence their subsequent reports in the face of countersuggestions, a number of the children in the Classmate conditions (\( n = 6 \)) responded to the countersuggestion questions by nominating
classroom interactions as the source of their memory for the target events: “Brianna told me,” “I heard it from the kids,” and “I don’t know. I know it. Everybody knows it.”

Responses to peer conformity questions

Next we sought to examine the extent to which the children who did not report one or both of the target activities in response to open-ended or specific questions (i.e., initial denials) later acquiesced when confronted with peer conformity questions. The rate of change from initial denials of target activities to subsequent assents was calculated for each experimental condition and subjected to ANOVA. Table 3 lists the number of the children’s responses that were initial denials, the number of subsequent assents in response to peer conformity probes, and the rate of change from initial denials to assents following peer conformity questions (number of subsequent assents/number of initial denials) for the children in each experimental condition. Thus, the maximum value for the first column ranged from 32 to 36 (the number of children in each experimental condition times two target activities). As can be seen in the far right column on Table 3, the level of participation affected the children’s rates of change from initial denials of target activities to subsequent assents in response to peer conformity questions, $F(2, 59) = 4.19, p < .05$. Specifically, the rate of change for the children in the Classmate conditions ($M = .86$) was considerably higher than for those in the Control conditions ($M = .47$), $F(1, 59) = 8.38, p < .01$. The rates of change for the Witness ($M = .56$) and Classmate conditions did not differ significantly, $p = .12$. Note that this ANOVA included only those children who had not already reported both of the target activities in response to open-ended or specific probes.

Further illustrating the participation condition effect on the children’s responses to peer conformity questions, the rates of change for the children in the Witness and Control conditions did not differ significantly from the .50 value that would be expected on the basis of chance. In contrast, the rate of

<table>
<thead>
<tr>
<th>Experimental Group</th>
<th>Initial denials</th>
<th>Assents</th>
<th>Rate of change</th>
</tr>
</thead>
<tbody>
<tr>
<td>Witness/Neutral Interview</td>
<td>6</td>
<td>3</td>
<td>.50 (.55)</td>
</tr>
<tr>
<td>Classmate/Neutral Interview</td>
<td>18</td>
<td>15</td>
<td>.83 (.38)</td>
</tr>
<tr>
<td>Control/Neutral Interview</td>
<td>21</td>
<td>10</td>
<td>.48 (.51)</td>
</tr>
<tr>
<td>Witness/Suggestive Interview</td>
<td>3</td>
<td>2</td>
<td>.67 (.57)</td>
</tr>
<tr>
<td>Classmate/Suggestive Interview</td>
<td>4</td>
<td>4</td>
<td>1.00 (.00)</td>
</tr>
<tr>
<td>Control/Suggestive Interview</td>
<td>13</td>
<td>6</td>
<td>.46 (.52)</td>
</tr>
</tbody>
</table>
change for those in the Classmate conditions was relatively high and exceeded chance expectation, indicating that these children were especially likely to acquiesce to peer conformity questions. Note, however, that the data from the Witness conditions should be interpreted cautiously because a relatively small number of these children initially denied the target activities and consequently were asked peer conformity questions.

### Narrative length of reported target events

It was of interest to explore group differences in the extent to which the children’s reports of target activities were accompanied by narrative detail because examination of the interview transcripts revealed that many of the Classmate/Neutral Interview children produced surprisingly elaborate narratives of these nonexperienced activities. To illustrate, consider the following excerpts: “He spilled his coffee on his self. . . and he burned his self and hadded to go to the hospital,” “He put it [the soiled treasure map] in the bathroom to dry,” “He said, ‘Oh dear, I spilled my coffee, I’m going to take it to wash it in the kitchen,’ and Maryanne said, ‘No you can’t,’” “He just went and walked away and then we just got in big trouble . . . all my friends and he had to be punished for a whole weekend,” “The ladies in the cafeteria cleaned it because he didn’t have a mop,” and “They took him away and put him in jail.”

To understand the extent to which the children’s reports of target activities were accompanied by narrative detail, the “utterance idea” coding scheme described earlier was used. Narrative length was coded for the subset of children who reported at least one target activity at the open-ended or specific level of questioning. Table 4 shows the average Witnessed-Suggested, Constructive, and Total (Witnessed-Suggested plus Constructive) utterance ideas by experimental group. As can be seen in the far right column displaying Total utterance ideas, the children’s reports of target activities contained a considerable amount of narrative detail. Most interesting, the Total number of utterance ideas for the Classmate ($M = 12.25$) and Control children ($M = 9.40$) was much greater compared to that of the Witness

### Table 4

Mean number utterance ideas (and standard deviations) reported at the final interview as a function of experimental group

<table>
<thead>
<tr>
<th>Experimental Group</th>
<th>Witnessed-Suggested</th>
<th>Constructive</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Witness/Neutral Interview ($n = 14$)</td>
<td>4.71 (2.09)</td>
<td>1.14 (1.23)</td>
<td>5.85 (2.98)</td>
</tr>
<tr>
<td>Classmate/Neutral Interview ($n = 6$)</td>
<td>2.33 (.52)</td>
<td>9.00 (2.10)</td>
<td>11.33 (2.34)</td>
</tr>
<tr>
<td>Control/Neutral Interview ($n = 0$)</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Witness/Suggestive Interview ($n = 16$)</td>
<td>5.56 (3.76)</td>
<td>2.44 (2.31)</td>
<td>8.00 (5.50)</td>
</tr>
<tr>
<td>Classmate/Suggestive Interview ($n = 14$)</td>
<td>2.71 (.99)</td>
<td>9.93 (4.01)</td>
<td>12.64 (4.68)</td>
</tr>
<tr>
<td>Control Suggestive Interview ($n = 10$)</td>
<td>3.00 (1.33)</td>
<td>6.40 (2.95)</td>
<td>9.40 (4.01)</td>
</tr>
</tbody>
</table>
children ($M = 7.00$), indicating that the children’s false narratives were more elaborate than the children’s true narratives. Also, the Classmate children’s narratives contained more detail than those in the Control condition, suggesting that exposure to peer witnesses can elevate the embellishment of reports of non-witnessed events above and beyond that of peer witnesses. Note, however, that the $n$s for the Classmate/Neutral Interview and Control/Neutral Interview groups were quite low. Because of this variation in the number of subjects contributing to the means for the Neutral Interview conditions, the differences in narratives of target activities among these children were not evaluated statistically. Nonetheless, a one-way ANOVA carried out on the Suggestive Interview groups confirmed a significant effect of participation condition, $F(2, 37) = 3.68, p < .05$. Follow-up contrasts indicated that the children in the Classmate/Suggestive Interview group reported a greater number of utterance ideas compared to those in the Witness/Suggestive Interview group, $F(1, 37) = 7.16, p < .01$, supporting the claim that the Classmate children described the target activities at greater length than the children who actually witnessed these activities.

Inspection of Table 4 also indicates group differences in the children’s production of Witness-Suggested versus Constructive utterance ideas. As can be seen in the first and second columns, the non-witness children were especially likely to generate constructive details, rather than use their peers’ experiences or the interviewer’s suggestions as the basis of their reports of target activities. In contrast, the Witness children were more inclined to limit their narratives to actually experienced or interviewer suggested details. A one-way ANOVA confirmed significant participation group differences among the Suggestive Interview children, $F(2, 37) = 21.90, p < .001$, with the children in the Classmate/Suggestive Interview condition reporting more Constructive utterances ideas than those in the Control/Suggestive Interview condition, who in turn reported more Constructive utterance ideas than those in the Witness/Suggestive Interview condition, $F_{s}(1, 37) \geq 7.56, ps \leq .01$. Thus, peer exposure induced the construction of original embellishments of nonexperienced activities above and beyond the effects of suggestive interviews.

**Discussion**

The results of this investigation show that naturally occurring conversations with peers can lead to false reports of nonexperienced events, even in the absence of exposure to misleading information from adults. The children in the Classmate conditions reported a greater proportion of target activities compared to those in the Control conditions, demonstrating that exposure to children who witnessed a salient event can lead non-participants to produce false reports of this event. Further, when exposed to suggestive
interviews, the total recall scores of the children in the Classmate conditions were elevated to the level of those in the Witness conditions, indicating that the combined effects of exposure to peer witnesses and suggestive interviews may be especially detrimental, obscuring or even eliminating differences in levels of recall between true and false reports.

Perhaps the most interesting result to emerge from this study is that the children in the Classmate groups were more likely than those in the Control groups to state that they actually saw the target activities (as opposed to merely having heard about them). These findings provide original evidence that conversations with agemates are capable of influencing not only children’s reports of their experiences, but also their underlying beliefs about what happened. However, the difference between the Classmate and Control conditions disappeared when the children were posed a countersuggestion, suggesting that the negative effects of overhearing nonoccurring information from peers may be limited to conditions under which children’s reports of witnessing an event are not challenged by an interviewer.

Further supporting the role of peers in children’s remembering, rates change from initial denials of target activities to subsequent assents following peer conformity questions varied as a function of participation level, with only the children in the Classmate conditions having higher than chance rates of changing an initial denial to an assent. Thus, these data suggest that exerting peer pressure may be particularly successful when children believe that their peers may have witnessed events that are consistent with information that is being suggested by an interviewer. As such, for some children, the peer conformity probes may have enabled access to information in memory that they heard about from peers but were unwilling or unable to report at the open-ended or specific level of questioning.

These findings are consistent with Bruck et al.’s (in press) demonstration that peer pressure can induce false reports, but previous work by Garven et al. (2000) has not revealed an effect of peer conformity on accuracy. Although the reasons for this discrepancy remain unclear, there are at least three possibilities. First, the participants in current study and in Bruck et al. (in press) were preschoolers, whereas Garven et al. (2000) examined 5- to 7-year-olds. Given preschoolers’ heightened susceptibility to various forms of suggestion (see Ceci & Bruck, 1995), one would expect them to be more vulnerable than older children to suggestions involving peers. Second, preschoolers have particular difficulty with yes/no questions (e.g., Gordon & Follmer, 1994), and consequently would be expected to have more trouble than older children responding accurately to yes/no queries implicating peers. Third, the suggestions in the current study were highly plausible with what might have happened during the archaeology event, and plausibility is well known to affect the ease with which study participants will accept misinformation (e.g., Pezdek, Finger, & Hodge, 1997). Additional research is needed to tease apart these alternatives.
A related finding is that exposure to peer witnesses promoted the embellishment of reports of nonexperienced events among non-witnesses, such that the children in the Classmate conditions provided more detailed narratives compared to those in the Control conditions. Further, these data are consistent with a growing body of work demonstrating that narrative length is not diagnostic of true versus false accounts (Bruck et al., in press; Lamb, 1998; Poole & Lindsay, 2001, 2002)—the false narratives of the children in the Classmate and Control conditions contained more details than the true narratives of those in the Witness conditions. Examination of the content of the children's narratives indicated that this group difference occurred primarily because the children who did not witness the target events were more likely than those who actually saw these events to generate constructive details when encouraged to discuss these events. Admittedly, however, the reports of those in the Witness groups were somewhat constrained by actually experiencing the target activities, whereas those in the Classmate and Control groups were not limited to reporting the details of an actual experience. Despite this constraint, as it exists in children's accounts elicited everyday and forensic settings, these findings demonstrate the potential dangers of encouraging children to elaborate following exposure to peer or adult suggestions.

Although the experimental procedures used in this study were not designed to test directly the mechanisms underlying the children's reporting errors, several pieces of evidence have implications for understanding the degree to which suggestibility is mediated by memory-source confusions as opposed to social factors (e.g., wanting to be part of the group that witnessed the target activities, or desiring to conform to suggestions by knowledgeable prior adult interviewers). First, many of the children's false narratives included a coherent blend of material gained from events they actually experienced intermixed with non-witnessed details adopted from interactions with peers or from intervening interviews. This finding is consistent with studies demonstrating that young children often confuse suggested information with aspects of experienced events, leading to difficulties differentiating between memories from different sources (Gopnik & Graf, 1998; Lindsay, Gonzales, & Eso, 1995; Poole & Lindsay, 2001, 2002). If true, then this would suggest that at least some of the children's false reports were due to actual memory confusions rather than socially driven acceptance of nonexperienced information. Along these same lines, the finding that the children in the Classmate conditions were significantly more likely than those in the Control conditions to state that they actually saw the target activities (as opposed to merely having heard about them) provides additional support that interactions with peers, in and of themselves, are capable of producing source confusions among children.

However, source confusions were likely only part of the problem. This is because some of the Classmate and Control children who recalled the target activities later indicated that they only heard about them. These children
may have initially reported the target activities because they trusted a peer’s or a prior interviewer’s version of what happened more than their own memory (McCloskey & Zaragoza, 1985), or because they misconstrued the intent of the interviewer’s questions as an assessment of both “remembered” and “known” information (as opposed to only remembered information) about the event in question (Tulving, 1985).

Taken together, these findings suggest that both memorial and social factors likely were at work, albeit to a different extent in each of the six experimental groups, thus supporting the idea that multiple pathways exist to false reporting. Nonetheless, further research clearly is needed to tease apart those situations that afford source confusions versus social acceptance of misinformation. Indeed, determining the boundary conditions of genuine memory change has considerable forensic relevance as no degree of supports for remembering could lead to the recovery of an original memory in situations where the initial representation has been replaced or altered by suggestions from others.

Conclusions

Perhaps the most important result to emerge from this study is the extent to which two very different sources of suggestibility, namely biased interviews and peer influence, induced spontaneous intrusions in the children’s accounts. Indeed, these data are consistent with an emerging body of work indicating that high rates of open-ended errors can arise from a range of suggestive adult-generated influences (Bruck et al., in press; Ceci et al., 1994a,b; Leichtman & Ceci, 1995), but go beyond existing studies by demonstrating that interactions with same-age peers can elevate spontaneous claims of false events even in the absence of exposure to misleading information from adults. As hinted above, these findings would seem to hold some applied relevance as they make clear that clinical and legal professionals must be cautious in using spontaneity as criteria for gauging the reliability of children’s reports, particularly in situations where opportunities to be influenced by peer witnesses may have taken place prior to the clinical or legal interview.

The current results also highlight the importance of considering contaminating factors outside of the formal interview environment in evaluations of the accuracy of children’s legal testimony. Although caution is required in generalizing these findings to situations involving child witnesses (the event was largely enjoyable, the children were not pressured, and there were no incentives to influence their reports), these data indicate that even in the absence of leading questions and under supportive interviewing conditions, children may nevertheless incorporate false information that emanated from peers rather than from interviewers. As such, peer generated suggestibility cannot be easily alleviated by the techniques currently used by forensic
experts to reduce bias in the way that interviewers question young children (see e.g., Poole & Lamb, 1998). Moreover, because the level of peer influence in this study may be pale compared to that which occurs in group therapy sessions or in joint police interviews, these data underscore the importance of considering the potential contaminating influence of peers in legal cases involving multiple child witnesses.

Acknowledgments

The research reported here was supported by Grant F32 MH12619 from the United States Public Health Service. The author gratefully acknowledges the cooperation of the children, parents, and teachers of the Early Childhood Center at Cornell, IACC Day Care Center, Ithaca Community Childcare Center, and Lansing Community Action Head Start, without whose interest and cooperation the study would not have been possible. Thanks are also due David Battin, Patricia Blumenauer, Katrin Mueller, Ashley Schumacher, and Lisa Tallon, for assistance in data collection and coding. Portions of this research were presented at the meeting of the Conference on Human Development, Charlotte, NC, April 2002.

References


