JOURNAL OF NARRATIVE AND LIFE HISTORY, *1*(1), 69–99
Copyright © 1991, Lawrence Erlbaum Associates, Inc.

Children's Testimony About a Stressful Event: Improving Children's Reports

Gail S. Goodman, Bette L. Bottoms, and
Beth M. Schwartz-Kenney
State University of New York at Buffalo

Leslie Rudy
Ohio State University

Age differences in children's ability to recount a stressful event were explored, as were several ways to improve children's reports. Seventy 3- to 7 year olds were videotaped while receiving inoculations at a medical clinic. It was predicted that multiple interviews would maintain memory and strengthen resistance to suggestion. It was also predicted that social support would ease intimidation and thus lessen children's suggestibility. To test these predictions, children were interviewed either once after a 4-week delay or twice, following 2- and 4-week delays, and under either "reinforcing" or "nonreinforcing" conditions. Age differences in answers to specific and misleading questions and in performance on a photo identification task were prevalent. However, multiple interviews and reinforcement supported more accurate reports. Training was effective in reducing false identifications on the photo identification task, especially for older children. Children's accuracy was unrelated to parental ratings of the stressfulness of the event. Our findings have implications for the testimony of child victim witnesses and for child–adult reconstruction of a child's past history. *(Psychology)*

How children and adults negotiate the reconstruction of a child's past is a question of critical importance to developmental psychology. In recent years, it has also become a critical question for the legal system. There are few times when the child–adult negotiation of the past is more consequential—or more controversial—than when children are interviewed in a forensic context or when they testify in courts of law.

Requests for reprints should be sent to Gail S. Goodman, Department of Psychology, State University of New York at Buffalo, Buffalo, NY 14260.

Concern about the accuracy of children's testimony results in large part from young children's more limited cognitive abilities compared to adults and from children's tendency, at least under certain conditions, to accept adults' rendition of past events. Indeed, a number of studies of children's reports of real-life events indicate that young children (e.g., 3 and 4 year olds) often recall less than older children and adults, tend to be more susceptible to suggestion, and are more likely to make false identifications in photo displays (e.g., Davies, Tarrant, & Flin, 1989; Goodman & Aman, in press; Goodman, Aman, & Hirschman, 1987; King & Yuille, 1987; but see Marin, Holmes, Guth, & Kovac, 1979; Zaragoza, 1987; Zaragoza & Wilson, 1989).

Although developmental differences in children's cognitive abilities clearly exist (Case, 1985; Fischer, 1980; Piaget, 1936/1952), children of a certain age or cognitive level have a range of abilities (Fischer & Bullock, 1984; Price & Goodman, 1990; Vygotsky, 1934/1978). Under supportive conditions (e.g., cognitive support provided by memory cues or practice, emotional support provided by warmth and affection), children may be able to perform at their "optimal level" (Fischer, 1980). It may thus be possible to develop techniques to optimize the accuracy of children's reports. For example, it may be possible to develop techniques for communicating to children that, instead of an interest in agreement, adults are more interested in what the child has to say about what did or did not happen.

In this study, we explore several techniques to improve young children's memory and reduce their suggestibility, that is, to optimize the child–adult negotiation of the past. To the extent that young children's testimony is deficient compared to that of older children and adults, researchers can provide a service by developing techniques to enhance children's ability to recount life events accurately. Moreover, the development of such techniques may provide insights into the causes of deficits in children's testimony.

In exploring techniques to improve children's reports, we examined children's memory and suggestibility for a stressful event, specifically, receiving inoculations at a medical clinic. Such events mimic, in important and ethical ways, features of victimization. Children often report experiencing stress and pain when they receive inoculations (Melamud, 1976; Steward & Steward, 1981). Although many children seem relatively undisturbed by the experience, other children cry, and some become nearly hysterical with fear. The more stressed children may fight back or try to flee and thus have to be restrained. Especially for these more highly stressed children, the experience is similar in important respects to an attack. Of course, one would not want to subject laboratory participants to such stressful conditions; however, it is possible to capitalize on naturally occurring stressful situations to study children's reports of stressful events.

We were particularly concerned with exploring children's reports in relation to a specific form of victimization—child sexual abuse. Children's and

adults' reports of childhood sexual assault have been viewed with skepticism for many years. For example, Freud (1905/1975) feared that such reports can result from sexual fantasy. More recently, concerns have focused on children's suggestibility.

Because of these concerns, interviewers' questioning in actual child sexual abuse investigations often becomes a key issue, particularly their use of leading questions. Legally, leading questions include such specific queries as "Did she put anything in your mouth" or "Did he touch your pee pee?" as well as more strongly worded questions such as "He touched your pee pee, didn't he?" (*Idaho v. Wright,* 1989; Myers, 1987). On the one hand, there may be a risk of obtaining false reports from children if leading questioning is used. On the other hand, because it can be difficult to obtain complete reports from children based on relatively vague free-recall questions ("What happened last week?" "Did anything ever happen to make you feel uncomfortable?"), interviewers may find it necessary to ask children more specific questions. In our study, we explored children's responses to leading questions about acts that might be of interest in child abuse investigations.

Former research along these lines indicates that, at least by the age of 4 to 5 years, children are surprisingly, though not completely, resistant to leading questions regarding acts related to child abuse, such as whether the children were hit, had their clothes removed, or had something put in their mouths (e.g., Goodman et al., 1987; Goodman, Hirschman, Hepps, & Rudy, in press; Rudy & Goodman, in press). In contrast, 3 year olds have been found to be more suggestible than older children in regard to such abuse-related questions (Goodman & Aman, in press).

What are the causes of young children's heightened suggestibility, and can we develop techniques to reduce the errors they make? Although several researchers have emphasized the role of memory in influencing children's suggestibility (e.g., Brainerd & Reyna, 1988; Goodman & Reed, 1986), socioemotional factors may also play an important role. In particular, the amount of intimidation versus comfort children perceive in a situation may influence their reports. In past research, we noticed, for example, that 3 year olds often seemed shy when interacting with an unfamiliar adult (Goodman & Reed, 1986). Intimidation may in part underlie increased suggestibility, including suggestibility on abuse-related questions. Evidence exists that interviewer status can affect suggestibility in adults (Loftus, 1979) and children (Ceci, Ross, & Toglia, 1987; but see Brigham, Van Verst, & Bothwell, 1986): Increased status is associated with increased suggestibility. In contrast, social support, such as the presence of a child's friend during an interview, has been found to increase children's accuracy (Moston, 1987). It is important, then, to investigate methods to reduce the intimidation that can be inherent in adults' questioning of children, thereby hopefully reducing young children's suggestibility.

In our study, we investigate a form of social support that we refer to as "reinforcement." Consistent with a behaviorist orientation, it has been proposed that interviewer practices, such as complimenting children on their performance, may reinforce inaccuracies and lead to false reports of abuse (e.g., Underwager & Wakefield, 1990). Thus, although such practices are often used by clinicians and developmental psychologists to establish rapport with children and to build interest and encouragement into interview and experimental situations, criticism can result when an investigator or therapist uses such techniques in a forensic interview with alleged child victims. The effects of such practices on the accuracy and completeness of children's reports require careful examination. Therefore, we decided to take a beginning step in examining the effects of reinforcement on children's testimony. For present purposes, we operationally defined reinforcement as an interviewer being particularly nice and comforting to children (e.g., complimenting the children on their performance, serving cookies and juice).

We were also interested in the effects of repeated interviews on children's accuracy. In actual child-abuse investigations, children are often repeatedly interviewed. Concerns have been raised that suggestive questioning used in one interview may contaminate children's reports in subsequent interviews. Although this remains a realistic possibility, memory development researchers generally find that repeated interviews help maintain children's memory (Baker-Ward, Hess, & Flannagan, 1990; Brainerd & Ornstein, in press; Brainerd, Reyna, Howe, & Kingma, 1990; Dent & Stephenson, 1979; Fivush & Hammond, 1989; Tucker, Mertin, & Luszcz, 1990).

However, these studies have generally concerned free recall, that is, answers to open-ended questions (e.g., "What happened?") or answers to specific but not intentionally misleading questions. We know of no studies that have examined the effects of repeated interviewing versus a single delayed interview on children's responses to deliberately misleading questions. If repeated interviewing heightens children's suggestibility, multiple interviews of children, especially employing misleading questions, might be detrimental to accurate reporting. But if repeated interviewing maintains accurate memory, and if memory provides a basis for resistance to suggestion (e.g., Goodman & Reed, 1986; Loftus, 1979), repeated interviewing may also maintain resistance to misleading questions and thus improve the quality of children's reports.

The experimental design often employed in research on the effects of a repeated versus a single delayed memory assessment is to test one group of children repeatedly, whereas another group is tested only once after a delay. This design permits examination of the effects of delay interval as well as repeated testing on the accuracy of children's report. Evidence now exists that developmental differences occur in forgetting rates (Brainerd et al., 1990). Thus, it is possible that after a delay, developmental differences will arise in the frequency and type of errors children make. Because of the potentially serious legal consequences of commission errors (errors in which the child

reports, either spontaneously or through simple affirmation, that something occurred that did not), we were particularly interested in examining whether older or younger children are more likely to make errors of commission after a delay.

The last improvement technique we examined concerns children's tendency to make false identifications of strangers in photo identification tasks. In a study that examined children's memory for a staged event, King and Yuille (1987) found that 8 to 11 year olds made false identifications on a target-absent lineup task 74% of the time, even when instructed that the target person might not be present. (A target-absent or "blank" lineup refers to a lineup in which the person of interest is not pictured.) Parker and Carranza (1989) also found that children make a relatively high number of false identifications on blank lineups. Both sets of researchers concluded that for children, photo lineups are inherently suggestive. Children may not be able to understand or deal effectively with the situational pressures, sometimes subtle, inherent in a photo lineup task: Even with a warning, the task may imply to children that making any choice is the best alternative.

Assuming, then, that lineups might be inherently suggestive to children, it is important to develop techniques to enhance children's performance on photo identification tasks. Children may not understand, even if told outright, that the person of interest might not be pictured. Thus, we expected that training on sample lineups would be beneficial. We developed several training trials in an attempt to communicate to children that the person of interest might not be pictured and that a correct answer might be not to point to any of the pictures.

Finally, given our interest in children's memory for stressful events, we explored the effects of stress on children's testimony. There are reasons to expect that stress might have positive effects on memory, but there are also reasons to expect the opposite. Some writers have emphasized the adaptive value of retaining memories of aversive experiences (Brown & Kulik, 1977; Goodman, Rudy, Bottoms, & Aman, 1990; Pillemer, 1984; but see Bohannon, 1988; McCloskey, Wible, & Cohen, 1988; Neisser, 1982), and physiological evidence indicates that stressful events are retained particularly well (Gold, 1987; Kety, 1970; McGaugh, 1989). Other writers have relied on the Yerkes–Dodson law (Yerkes & Dodson, 1908) and Easterbrook's (1959) expansion of it as a theoretical framework for predicting adverse effects of stress on memory (e.g., Deffenbacher, 1983; Loftus, 1979; Peters, 1987). According to the Yerkes–Dodson law, too little or too much stress is associated with decreases in memory, whereas a moderate level of stress is associated with a relative increase in memory.

Perhaps these two opposing views can be resolved if excessive stress has both positive and negative effects on memory (e.g., Baddeley, 1982; Christianson & Loftus, 1987; Goodman et al., in press; Heuer & Reisberg, 1990). For example, stress may narrow attention, leading to poor memory for certain

details, but might aid consolidation so that details that have been attended to are retained particularly well (Baddeley, 1982).

Research findings on the relation between stress and memory have been mixed. Although the majority of studies examining the effects of stress on children's reports indicate either positive or null effects (Goodman et al., in press; Oschner & Zaragoza, 1988; Steward, 1989; Warren-Leubecker, Bradley, & Hinton, 1988), others may have uncovered negative effects of stress on children's suggestibility and photo identifications (Peters, 1989; but see Goodman, in press). Mixed findings also characterize relevant research with adults (e.g., Bohannon, 1988; Brigham, Mass, Martinez, & Whittenberger, 1983; Harsch & Neisser, 1989; Heuer & Reisberg, 1990; Loftus & Burns, 1982; McCloskey et al., 1988; Yuille & Cutshall, 1986; for reviews, see Deffenbacher, 1983; Goodman & Hahn, 1987).

In a study quite similar to the current one, Goodman et al. (in press, Experiment 2) found that children who were the most severely stressed by receiving an inoculation—that is, children who cried and screamed, struggled to get away, and became nearly hysterical with fear—later recalled more about the event and were more resistant to suggestion than children who appeared to be less stressed. However, these findings were based on a relatively small number of children, and it remains to be seen if the same results will be found for other children. In other similar studies, no significant effects of stress were uncovered (Goodman et al., in press).

Former studies of the effects of stress on children's memory have not typically taken into account the possible effects of repeated interviewing on children's ability to report a stressful event. It is possible that repeated interviewing will result in hypermnesia (better memory over time) especially for events that are stressful (Erdelyi, 1985). It is also possible, however, that children may not be willing to recall over and over again the details of a stressful event and may thus recall less and less with repeated interviewing (Flin, in press).

In our study, 3- to 7-year-old children were videotaped while receiving oral polio vaccines and inoculations at a medical clinic. Parents rated the children's stress at the time of the inoculation. One group of children was interviewed about the event twice, once at 2 weeks and again at 4 weeks, and another group was interviewed only once at a 4-week delay. In this way, we could determine if repeated interviewing is detrimental to children's memory and suggestibility. Approximately half the children in each of these groups experienced interviews in a reinforcement condition, whereas the other half did not. Moreover, approximately half the children received practice on a series of photo identification tasks.

Several predictions were made. First, it was predicted that, compared to a delayed interview, repeated interviewing would help maintain children's memory and would also maintain resistance to suggestion. This prediction was based on the growing body of data indicating that repeated interviewing,

compared to delayed interviewing, results in more accurate reports. The extent to which a witness is susceptible to suggestion is often inversely related to the amount of information retained about an event (Goodman & Reed, 1986; Loftus, 1979; Warren, Hulse-Trotter, & Tubbs, in press).

Second, it was predicted that reinforcement would be associated with more accurate reporting in children, particularly in regard to their suggestibility. As explained earlier, lowering the intimidation value of the interviewer would be expected to heighten children's resistance to misleading suggestion.

Third, we predicted that practice with target-present and target-absent photo lineups would help children better understand that it is not necessary to point to someone's picture when presented with a photo lineup.

Fourth, it was predicted that stress would increase the amount of information children recalled about the event as well as their resistance to suggestion. This prediction was based on current empirical findings in the child witness literature as well as physiological evidence, as cited earlier.

METHOD

Participants

Seventy children participated, including 10 male and 13 female 3 and 4 year olds (M = 4 years, 3 months; range = 3 years, 4 months to 4 years, 11 months) and 21 male and 26 female 5 to 7 year olds (M = 5 years, 7 months; range = 5 years, 1 month to 7 years, 2 months). Participants came from families of lower to lower middle socioeconomic levels. Twenty-nine White, 25 Hispanic, and 16 Black children were included. Ten of the children were known to have visited the clinic once before. Children were randomly assigned to groups with the exceptions that an approximately equal number of boys and girls were assigned to the experimental conditions and that, because we were approaching people by surprise in the clinic waiting room, it was necessary to make some accommodations to the families' schedules.

None of the children were known to be ill. Families came to the clinic primarily because their children needed inoculations to enter public schools. Others came to maintain the standard schedule of shots recommended for children or, in a few cases, because shots were required to visit foreign countries. All the children were fluent in English.

Materials

Stress Rating Scale. A 6-point rating scale, ranging from *extremely happy or relaxed* (1) to *extremely frightened or upset* (6), was used by parents to indicate their perception of their child's stress level during the inoculation

procedure. These ratings served as the basis for assigning children to stress groups for purposes of statistical analysis.

Questionnaire. The interview consisted of a free-recall question followed by 35 *general* and 11 *abuse* questions about the visit to the inoculation clinic. There were three categories of general questions: (a) *person* questions (7 specific, 4 misleading), which pertained to the nurse's physical appearance; (b) *room* questions (7 specific, 4 misleading), which pertained to the appearance and contents of the room in which the inoculation was given; and (c) *action* questions (8 specific, 5 misleading), which pertained to the activities that occurred during the inoculation procedure and attributes of the child related to the activities. The abuse questions (4 specific, 7 misleading) concerned information that might be of particular interest in an actual child-abuse investigation, such as, "How many times did she hit you?"[1]

Specific questions were aimed at obtaining more detailed information from children than that typically elicited through free recall. Although we were not trying to mislead the children through the use of these questions, in many legal circles our specific questions would be considered leading. In contrast, by asking the misleading questions, we intentionally implied misinformation; for example, that other children said the nurse told them to keep a secret. The misleading questions often included "tag" phrases such as "didn't she?"

Questions were approximately balanced in terms of requiring *yes* or *no* answers for a correct response, but several of the questions required a more specific answer.

Photo identification task. A "blank" photo lineup was constructed for each of the six nurses who administered the inoculations. Each lineup included six women who matched the general description of the nurse. The nurse who actually gave the child the shot was not pictured. In addition, three training lineups were constructed: (a) a target-present lineup of six familiar animals; (b) a blank lineup consisting of pictures of six women who were ethnically similar to the child's mother; and (c) a six-person target-present lineup that included the research assistant who administered the training lineups. The animal and research assistant lineups were included to indicate that the target item might be present. The mother lineup was included to indicate that the target item might not be present.

[1]Two of the questions ("She didn't wear a white coat, right?" and "Did she take a picture of you?") were not scored, leaving 10 abuse questions and 3 misleading person questions. The former, Question 9, was not scored because the children's answers were often ambiguous. The latter, Question 38, was not scored because a few of the children inadvertently saw the video equipment and monitor and may have realized that they were being filmed.

Digit-span test. A digit-span test, taken from the McCarthy (1972) scales, was included to test for possible differences in basic memory ability among children in the different experimental and stress groups.

Achenbach Child Behavior Checklist (CBCL). The CBCL was included in this study to assess individual differences in behavioral adjustment among children in the different experimental and stress groups. The CBCL assesses internalizing (e.g., depression, somatic complaints) and externalizing (e.g., delinquency, aggressiveness) behavior problems through parental report. Its psychometric properties are commendable (see Achenbach & Edelbrock, 1983; Wolfe & Mosk, 1983).

Design

The design of the study was a 2 (Age) × 2 (Number of Interviews) × 2 (Reinforcement) factorial design. The number-of-interviews variable refers to the fact that children were given either one interview at a 4-week delay or two interviews at 2- and 4-week delays. This interview schedule generated several possible comparisons of the children's performance (see Table 1). To examine the effects of one versus two interviews when time since the inoculation was controlled, the performance of all children at 4 weeks could be examined with number of interviews as a between-subjects variable (i.e., the 4-week performance of children who had two interviews compared with the 4-week performance of children who had only one interview). To investigate changes in children's testimony on a first versus second interview, performance at 2 weeks versus 4 weeks could be compared for children who had two interviews, with time of interview varied within subjects. To examine the effects of delay (i.e., a 2-week delay vs. a 4-week delay), the children's performance at the first interview could be treated as a between-subjects factor.

TABLE 1
Experimental Design

	Interview Schedule	
Single Interview Group		4-Week Interview[a]
Repeated Interview Group	2-Week Interview[b]	4-Week Interview[c]

Note. Superscripts provide information regarding statistical comparisons. Comparison *a* versus *c* yields effects of a single interview compared to multiple interviews. Comparison *b* versus *c* examines effects of multiple interviews when initial questioning is followed by subsequent questioning. Comparison *a* versus *b* concerns effects of delay.

Procedure[2]

Phase 1. Participants were recruited at a medical clinic on the day they were to receive their inoculations. Parents were approached in the waiting room and asked if they would like to participate in research concerning children's reactions to medical procedures. Our interest in memory was not mentioned. When parents responded positively, the children were screened to ensure that they did not have any known perceptual or learning problems and were not scheduled to visit the clinic, a dentist's office, or another doctor's office in the following 4 weeks. This screening resulted in the elimination of only a few children from the study.

When children were called into the inoculation room by the nurse, surreptitious videotaping began. Most (60) of the children were first given an oral polio vaccine. The children then received typically one but as many as three shots, usually in the arm but at times in the thigh. The inoculations mostly consisted of shots for diphtheria, tetanus and pertussis, as well as measles, mumps, and rubella. After the inoculation(s), parents rated their children's inoculation stress level using the 6-point stress rating scale described earlier.

Phase 2. Children were questioned either once, at a 4-week delay, or twice, at both 2- and 4-week delays. Children were interviewed individually in a laboratory room or in their own homes, at the convenience of the family.

At the first interview, primary caretakers completed the CBCL by indicating the frequency with which their child displayed a standardized list of behaviors within the preceding 2-week period. Then all children received the digit-span test, interview questions, and photo identification task. Approximately half the children in each age group and each number-of-interviews condition received reinforcement during their interview, whereas the other half did not. Reinforcement consisted of beginning the interview with cookies and juice and of the interviewer being warm and friendly, smiling a lot, and giving the child considerable praise such as "You're doing a great job" or "You've got a great memory." These reinforcements were given randomly and regardless of accuracy. The other half of the children were not treated as warmly. They were not given cookies and juice, and the interviewer was more distant. The interviewer made comments such as "OK" or "all right" when necessary, but no comments regarding the children's accuracy, and she smiled only infrequently.

All children were shown the six-person blank lineup for the nurse and asked

[2]Our procedures were approved by the University's Human Subjects Review Committee. Note that the inoculations were not imposed on the children by us, but were rather given as part of the children's standard medical care. Furthermore, parents were fully informed about the purpose of our study at the start of Session 2 before testing began. At that time, they were shown the questionnaire and requested to cross off any questions they did not want asked. Parents rarely crossed off any questions.

if the nurse who gave them the shot was pictured. The children were cautioned that the nurse might not be in the lineup and to point to a picture only if they were sure it was a picture of the nurse who gave them the inoculation. Half the children experienced three practice trials before they attempted the lineup. For the first practice trial, the child was shown pictures of familiar animals and asked to point to an animal (e.g., a bear) that was present. In the second practice lineup, the child was shown a picture of the interviewer embedded in pictures of women similar in appearance to the interviewer. The child was instructed by the interviewer that the interviewer's picture might or might not be included and to point to a picture only if the child was sure that it was a picture of the interviewer. In the third practice lineup, the child was shown pictures of women who were ethnically similar to the child's mother. The child was told that a picture of her mother might or might not be present in the lineup and to point to a picture only if the child was sure that the mother's picture was present. To ensure that effects of practice could be properly inferred, the order of the latter two lineups varied across children. In this way, the children could not be correct by simply guessing based on a predictable pattern set by the order in which the practice lineups were presented. If mistakes were made on the practice trials, the children were immediately corrected, and the interviewer explained why the choice was wrong.

For children who received multiple interviews, the entire Phase 2 procedure was repeated at the 4-week delay test, including the training procedure. Thus, children maintained their status in the reinforcement and practice conditions across interviews.

RESULTS

Preliminary analyses revealed few if any effects associated with sex, ethnicity, or interview location, and sporadic effects of reinforcement. Therefore, the data were collapsed across these factors, except where noted. Preliminary analyses also indicated that although children who had only one interview did not differ from those who had two interviews in digit-span scores, they did experience significantly greater exposure time to the nurse, $F(1, 65) = 5.14$, $p < .05$, and had significantly higher CBCL scores, $F(1, 57) = 4.77, p < .05$. These two variables were thus used as covariates in analyses concerning number of interviews. The effects of number of interviews to be reported maintained significance even when CBCL and exposure time to the nurse were statistically controlled. The effects to be reported also maintained significance or closely approached significance when the number of previous visits to the clinic was used as a covariate.

A series of multivariate analyses of variance (MANOVAs) was conducted for the general questions, with proportion of correct responses, omission errors, or commission errors to the person, room, and action questions as depen-

dent measures. A series of analyses of variance (ANOVAs) was employed to analyze: (a) the number of correct or incorrect units of information provided in free recall; (b) the proportion of correct responses, omission errors, or commission errors to the abuse questions; and (c) correct versus incorrect responses on the photo identification task.

Effects of Repeated Interviewing

The effects of repeated interviewing were examined in two ways. For comparisons involving 4-week interviews for children who had one versus two interviews, 2 (Age) $\times$ 2 (Number of Interviews) analyses were conducted, with both factors varying between subjects. For comparisons involving children interviewed twice, 2 (Age) $\times$ 2 (Time of Interview) analyses were conducted, with time of interview varying within subjects. Because these latter analyses largely duplicate the first, significant findings from them are presented only when they provide additional insights into the effects of repeated interviewing.

Recall

Recall protocols were scored for correct and incorrect units of information. For example, if a child who received a shot in the arm said, "*She poked a needle in my leg,*" the statement would be scored as 3 correct units (1 for indicating being "*poked,*" 1 for indicating the needle as the instrument of poking, and 1 for indicating that a woman did the poking) and 1 incorrect unit (for indicating being poked in the child's leg rather than the child's arm). Only information that was directly verifiable from the videotape of the inoculation or from the layout of the clinic was scored. Two independent raters scored 20% of the protocols. Proportion of agreement was .75, indicating acceptable reliability. Disagreements were resolved by discussion. The remaining protocols were scored by one of the two raters. Mean scores are presented in Table 2.

Separate ANOVAs considering the amount of correct and incorrect information recalled at the 4-week interviews revealed no significant main effects or interactions associated with number of interviews or age. However, age and reinforcement significantly influenced the amount of incorrect information recalled. Specifically, older children gave fewer inaccurate statements (M = 0.42) than younger children (M = 1.68) and children who received reinforcement made fewer incorrect statements (M = 0.55) than children who did not receive reinforcement (M = 1.09), $F(1, 66) = 5.14$, $p < .05$.

Person, Room, and Action Questions

Specific questions. One goal of this study was to examine how multiple interviews, compared to a single delayed interview, affect children's accuracy in answering specific questions. As can be seen in Table 3, children interviewed

TABLE 2
Recall Scores as a Function of Interview Schedule and Age

	2 Weeks		4 Weeks (Repeat)		4 Weeks (Delay)	
Response Type	*Younger*	*Older*	*Younger*	*Older*	*Younger*	*Older*
	(10)	(21)	(10)	(21)	(13)	(26)
Correct[a]	4.00	4.00	4.78	4.27	3.00	3.50
Incorrect[b]	0.78	0.50	1.56	0.36	1.77	0.46

Note. Number in parentheses indicates number of children in each group. [a]For analyses concerning effects of repeated interviewing and delay, all $Fs(1, 66) \leq 2.81$, *ns.* [b]For analysis concerning effects of repeated interviewing, age effect, $F(1, 66) = 14.70, p < .001$. For analysis concerning delay, age effect, $F(1, 66) = 6.11, p < .05$.

TABLE 3
Proportion Scores on the Specific Questions as a Function of Interview Schedule, Age, and Type of Question

	2 Weeks		4 Weeks (Repeat)		4 Weeks (Delay)	
Response Type	*Younger*	*Older*	*Younger*	*Older*	*Younger*	*Older*
Correct[a]						
Person	.45	.52	.51	.56	.35	.46
Action	.61	.81	.64	.81	.64	.67
Room	.71	.72	.71	.77	.64	.64
Omission[b]						
Person	.05	.05	.07	.06	.07	.02
Action	.16	.11	.16	.13	.24	.13
Room	.11	.11	.11	.09	.13	.15
Commission[c]						
Person	.23	.26	.29	.19	.31	.25
Action	.22	.07	.19	.05	.12	.13
Room	.16	.16	.21	.10	.18	.16

[a]For univariate analyses concerning effects of repeated interviewing, person questions, $F(1, 66) = 6.55, p < .025$, and room questions, $F(1, 66) = 4.98, p < .05$; multivariate effect of repeated interviewing, $F(3, 64) = 4.16, p < .025$. For effects of age, univariate tests for action questions, $F(1, 66) = 7.64, p < .01$; multivariate effect of age, $F(3, 64) = 3.27, p < .05$. For analyses concerning delay, multivariate effect of age, $F(3, 64) = 3.18, p < .05$. [b]All multivariate $Fs(3, 64) \leq 2.19$, *ns.* [c]For analyses concerning effects of repeated interviewing, all $Fs(3, 64) \leq 2.23$, *ns.* For univariate analyses concerning delay, significant Age × Delay interaction for action questions only, $F(1, 66) = 5.08, p < .05$; multivariate Age × Delay interaction, $F(3, 64) = 3.09, p < .05$. Simple effects analyses revealed a significant age effect only for children at the 2-week delay, $F(1, 29) = 7.14, p < .05$.

twice rather than once were more accurate on specific person and room questions, but not on specific action questions. We were also interested in age differences. Although older children were significantly more accurate than younger children, univariate tests revealed that significant age differences occurred for specific action questions only.[3] Separate MANOVAs considering omission and commission errors revealed no significant effects associated with number of interviews or age.

In summary, children were more accurate in answering the specific questions when they were interviewed twice rather than once, and older children were more accurate than younger children, particularly on the action questions. However, there were no significant age differences in the types of errors made.

Misleading questions. We were also interested in examining the effects of repeated interviews on children's responses to the misleading questions, our main measure of suggestibility. For correct answers to these questions, multivariate effects associated with number of interviews were nonsignificant (see Table 4). However, when the within-subject factor of time of interview was considered for the subset of children interviewed twice, it was evident that, over time, children became less suggestible in response to the misleading person questions and room questions. The children's resistance to suggestions concerning action questions, which was relatively high, did not significantly change with repeated questioning.

Age differences were apparent, with older children more accurate than younger children in response to misleading action questions and misleading person questions, but not to misleading room questions. The age difference in response to misleading questions resulted in part from age differences in omission errors. Although number of interviews did not have a significant effect on omission errors, age did, with younger children making more omission errors than older children in response to action questions, but not in response to person or room questions.

However, in a 2 (Age) $\times$ 2 (Reinforcement) MANOVA concerning omission errors to the person, action, and room questions, the significant main effect of age comparable to that already reported was subsumed by a significant Age $\times$ Reinforcement interaction, $F(3, 64) = 3.07, p < .05$, specifically with regard to room questions, $F(1, 66) = 4.04, p < .05$. Simple effects analyses

[3]It might be argued that both correct and "don't know" responses should be combined for analysis, because together they index the proportion of information provided by the children that was not inaccurate. For this data set, results of analyses concerning correct plus "don't know" responses were virtually identical to those concerning correct responses only. However, mean scores, especially to the person questions, were often raised when "don't know" responses were included.

TABLE 4
Proportion Scores on the Misleading Questions as a Function of Interview Schedule, Age, and Type of Question

	2 Weeks		*4 Weeks (Repeat)*		*4 Weeks (Delay)*	
Response Type	*Younger*	*Older*	*Younger*	*Older*	*Younger*	*Older*
Correct[a]						
Person	.42	.59	.67	.75	.41	.59
Action	.75	.71	.61	.78	.52	.74
Room	.33	.32	.42	.41	.25	.34
Omission[b]						
Person	.06	.09	.03	.07	.15	.10
Action	.19	.25	.39	.23	.46	.21
Room	.06	.13	.14	.16	.15	.13
Commission[c]						
Person	.50	.25	.31	.17	.40	.24
Action	.00	.00	.00	.00	.00	.04
Room	.44	.37	.44	.29	.48	.36

[a]For univariate analyses concerning effects of repeated interviews, significant age effects for action questions, $F(1, 64) = 7.45$, $p < .05$, and person questions, $F(1, 66) = 4.88$, $p < .05$; multivariate age effect, $F(3, 64) = 3.25$, $p < .05$. For univariate analyses concerning effects of time of interview, significant time of interview effect for person questions, $F(1, 29) = 11.20$, $p < .01$, and room questions, $F(1, 29) = 6.62$, $p < .025$; multivariate effect of time of interview, $F(3, 27) = 5.77$, $p < .01$. [b]For univariate analyses concerning effects of repeated interviews, age effect for action questions, $F(1, 66) = 8.42$, $p < .01$; multivariate effect, $F(3, 64) = 3.87$, $p < .025$. For univariate analyses concerning delay, Age × Delay interaction for action questions, $F(1, 66) = 6.54$, $p < .05$; multivariate Age × Delay interaction, $F(3, 64) = 4.51$, $p < .01$. Simple effects analyses revealed a significant age effect for children first interviewed at 4 weeks, $F(1, 37) = 15.21$, $p < .001$, and a significant delay effect for younger children, $F(1, 20) = 7.96$, $p < .05$. [c]For multivariate analyses concerning effects of repeated interviews, all Fs$(3, 64) \leq 2.48$, *ns*. For multivariate analyses concerning effects of delay, age effect, $F(3, 64) = 3.17$, $p < .05$.

revealed that younger children who received reinforcement made more omission errors to the room questions ($M = .21$) than younger children who did not receive reinforcement ($M = .08$), $F(1, 20) = 8.24$, $p < .01$. No other simple effect involving this interaction was significant, and the main effect of reinforcement also failed to reach significance, $F(3, 64) = 1.50$.

The proportion of commission errors was not significantly related to number of interviews, age, reinforcement, or the interaction of these factors. As can be seen in Table 4, children rarely made any commission errors to the action questions.

Examination of the data indicated that the effects of reinforcement were particularly prominent for the children in the single-interview condition. These children's memory should have been the weakest because of the longer delay before their first interview. A 2 (Age) × 2 (Reinforcement) MANOVA with

proportion of commission errors on the person and room questions as dependent measures was performed including children in the single-interview condition only. None of the children made a single commission error to the misleading action questions, so these were not included in the analysis. The multivariate Age × Reinforcement interaction was significant, $F(2, 34) = 3.44$, $p < .05$. Univariate tests indicated that the interaction was significant for both the person questions, $F(1, 35) = 7.01$, $p < .025$, and the room questions, $F(1, 35) = 4.32$, $p < .05$. Simple effects analyses revealed that children did not differ significantly by age in the proportion of commission errors made to the person and room questions, respectively, when they were in the reinforcement condition (younger Ms = .20 and .36; older Ms = .27 and .41). However, significant age differences occurred for children in the no-reinforcement condition. Specifically, in response to the person and room questions, younger children (Ms = .63 and .63) made significantly more commission errors than did older children (Ms = .21 and .31), $F(1, 18) = 11.88$, $p < .01$, and $F(1, 18) = 7.74$, $p < .025$, respectively.

To summarize, children in the repeated-interview and single-interview conditions did not significantly differ in their suggestibility. However, children interviewed twice were less suggestible the second, compared to the first, time they were interviewed. Also, older, compared to younger, children answered the misleading questions with greater accuracy. The significant age difference resulted in part from age differences in the omission error rate.

However, reinforcement as well as age affected the types of errors made, although significant reinforcement effects on commission errors were only apparent when the single-interview children's data were considered. In that case, age differences in commission errors occurred only for children in the no-reinforcement condition. In contrast, when children were interviewed with reinforcement, significant age differences in commission errors disappeared. Nevertheless, the effects of reinforcement were not all positive: Younger children made more omission errors to the room questions if they were interviewed in the reinforcement condition than if interviewed in the no-reinforcement condition.

Abuse Questions

Because both the specific and misleading abuse questions would be considered leading in legal circles and because relatively few specific abuse questions were asked, we grouped the specific and misleading abuse questions for purposes of statistical analysis. An analysis of children's accuracy in response to abuse questions revealed a significant main effect of number of interviews (see Table 5). Rather than children becoming more likely to make errors when asked the abuse questions twice, children who had two interviews were more accurate the second time these questions were asked than were children who

TABLE 5
Proportion Scores on the "Abuse" Questions as a Function of Interview Schedule and Age

	2 Weeks		*4 Weeks (Repeat)*		*4 Weeks (Delay)*	
Response Type	*Younger*	*Older*	*Younger*	*Older*	*Younger*	*Older*
Correct[a]	.60	.81	.66	.82	.62	.70
Omission[b]	.24	.13	.19	.12	.19	.16
Commission[c]	.17	.05	.15	.05	.18	.11

[a]For analysis concerning effects of repeated interviews, effect of number of interviews, $F(1, 66) = 4.27$, $p < .05$, and age, $F(1, 66) = 8.80$, $p < .01$. For analysis concerning delay, age effect, $F(1, 66) = 11.68$, $p < .01$. [b]For analysis concerning effects of repeated interviews, age effect, $F(1, 66) = 4.20$, $p < .05$, and for analysis concerning delay, age effect, $F(1, 66) = 8.59$, $p < .01$. [c]For analysis concerning effects of repeated interviews, age effect, $F(1, 66) = 7.96$, $p < .01$, and for analysis concerning delay, age effect, $F(1, 66) = 6.82$, $p < .05$.

had only one interview. Also, older children were more accurate than younger children.

There were no significant main effects or interactions associated with number of interviews in the proportion of omission or commission errors made by the children. Older children made fewer omission and commission errors than younger children regardless of the number of times they were interviewed.

Interestingly, when a 2 (Age) $\times$ 2 (Reinforcement) ANOVA was conducted on commission errors, the main effect of reinforcement approached significance, $F(1, 66) = 3.47$, $p = .07$, as did the Age $\times$ Reinforcement interaction, $F(1, 66) = 3.16$, $p = .08$. Moreover, when the children's sex was entered as an additional factor, the Age $\times$ Reinforcement interaction was significant, $F(1, 62) = 4.20$, $p < .05$, whereas sex effects were not. Analyses of simple effects were conducted. In the no-reinforcement condition, younger children ($M = .24$) made significantly more commission errors than older children ($M = .08$), $F(1, 32) = 14.03$, $p < .01$. However, in the reinforcement condition, there was no significant difference between the younger ($M = .12$) and older ($M = .09$) children's performance, $F(1, 34) = 0.49$.

Further analyses revealed that the Age $\times$ Reinforcement interaction was again especially evident for the children who had only one interview at 4 weeks. ANOVA with age and reinforcement as between-subjects factors for the children who had only one interview revealed a significant Age $\times$ Reinforcement interaction, $F(1, 35) = 5.04$, $p < .05$. Simple effects analyses indicated that, among children who did not receive reinforcement, older children ($M = .09$) made fewer commission errors than younger children ($M = .28$), $F(1, 17) = 9.18$, $p < .025$. Older children ($M = .13$) and younger children ($M = .10$) who received reinforcement did not differ significantly in

their commission error rate. For the younger children, the effect of reinforcement approached but did not reach significance, $F(1, 11) = 4.08$, $p < .10$.

Because commission errors to the abuse questions have considerable legal importance, a closer examination of these errors was made. Most of the commission errors to the abuse questions consisted of the children simply saying *yes* or *no* without embellishment. As can be seen in Table 6, children made the most errors in response to the question about whether the nurse told them to keep a secret. This question was particularly leading in that it suggested that the false information was true for all the other children as well as the child being interviewed. The question that produced the next greatest number of errors concerned whether the children had been hit. It became apparent during the questioning, however, that many of the children interpreted our question as if receiving the shot meant being hit. When asked how many times they had been hit these children stated the number of shots they had received and pointed to the places on their bodies where the shots had been given. Also, some children seemed to interpret the question about tickling the nurse as if we had asked if the shot tickled. For example, one child responded, "*Yes, it tickled.*" Other children made errors that seemed to involve prototypical expectations. One, for example, claimed that the nurse put a stick in her mouth (a thermometer or tongue depressor?) when in fact the nurse had not. Yet other errors might have resulted from previous visits to the doctor's office; for example, one child claimed that her mother took the child's pants down to get the shot when in fact the mother had not done so during the doctor's office visit of interest. The children never provided spontaneous information or embellishments concerning genital contact.

Overall, children who were interviewed multiple times were more, not less, accurate in response to the abuse questions. Significant age differences existed in the children's ability to answer the abuse questions correctly and in the types of errors made. However, reinforcement modulated whether age differences in commission errors occurred. When children were interviewed in the reinforcement condition, significant age differences in commission error rate did not appear.

Photo Identification

Children's performance on the photo identification task at the 4-week interviews was considered next. Because the children responded to a "blank" lineup, pointing to a picture was an incorrect response and saying the nurse was not pictured was a correct response. A 2 (Age) $\times$ 2 (Number of Interviews) between-subjects ANOVA revealed no significant difference in the number of false identifications made by children who had one interview ($M = .62$) versus children who had two interviews ($M = .69$). The Age $\times$ Number of Interviews interaction also failed to reach significance. We therefore collapsed

TABLE 6
Proportion of Commission Errors to the Abuse Questions as a Function of Age, Interview Schedule, and Reinforcement Condition

Question		*Younger Children*			*Older Children*		
		2 Weeks	*4 Weeks (Repeat)*	*4 Weeks (Delay)*	*2 Weeks*	*4 Weeks (Repeat)*	*4 Weeks (Delay)*
1. She didn't have any clothes on, did she?	R	.00	.00	.08	.00	.00	.08
	NR	.30	.10	.08	.04	.04	.04
2. That person didn't touch you, did she?	R	.00	.00	.00	.00	.00	.00
	NR	.00	.00	.00	.00	.00	.00
3. The nurse kissed you on the mouth, didn't she?	R	.00	.00	.00	.00	.00	.04
	NR	.00	.00	.00	.00	.00	.04
4. The person who gave you the shot made you take your clothes off, didn't she?	R	.00	.00	.00	.00	.00	.04
	NR	.00	.10	.08	.00	.00	.04
5. Did she put anything in your mouth?	R	.10	.10	.00	.00	.00	.00
	NR	.20	.20	.00	.00	.04	.04
6. She asked you to tickle her, didn't she?	R	.20	.20	.08	.04	.04	.08
	NR	.10	.40	.31	.04	.00	.08
7. How many times did she hit you?	R	.10	.10	.15	.04	.04	.12
	NR	.10	.10	.31	.08	.04	.08
8. Did she touch you anywhere other than your arm or leg?	R	.10	.00	.00	.00	.00	.04
	NR	.00	.00	.10	.04	.04	.00
9. Didn't she touch your bottom?	R	.10	.00	.00	.00	.00	.00
	NR	.00	.00	.15	.00	.00	.08
10. The nurse told all the other kids to keep a secret. Didn't she tell you to keep a secret, too?	R	.10	.30	.15	.04	.04	.15
	NR	.20	.20	.31	.04	.04	.15

Note. R = Reinforcement. NR = No reinforcement.

across the number-of-interviews variable to examine effects of practice on the children's false identification rate.

Practice led to improved performance on the photo identification task. A 2 (Age) $\times$ 2 (Practice) analysis revealed that children who received practice made fewer false identifications ($M = .51$) than those who did not receive practice ($M = .80$), $F(1, 66) = 5.90$, $p < .05$. Also, younger children were less accurate ($M = .80$) than older children ($M = .52$), $F(1, 66) = 5.76$, $p < .05$. The Age $\times$ Practice interaction was not significant; however, planned comparisons revealed that only the accuracy of the older children significantly improved with practice ($M = .33$) compared with no practice ($M = .70$), $F(1, 66) = 7.67$, $p < .01$. The younger children's performance remained fairly poor *with* ($M = .71$) or *without* ($M = .89$) the practice trials.

Thus, when presented with the photo identification task, children on their own made a high rate of false identifications of the nurse. However, practice on target-present and target-absent lineups significantly reduced older children's tendency to make false identifications.

Effects of Delay

We were also concerned with the effects of delay on the accuracy of the children's reports. For example, we wanted to explore whether, over time, children make more commission or omission errors and how the error rate might vary over time with age. A series of 2 (Age) $\times$ 2 (Delay) between-subjects analyses examined the effects of delay on younger and older children's first-interview performance. Because effects of age and reinforcement were generally consistent with those already reported, we concentrate in this section mainly on significant effects involving the delay factor.

Recall

The amount of information recalled by children interviewed after a 2-week delay was compared with the amount of information recalled by children whose first and only interview was after a 4-week delay. Surprisingly, the analyses indicated no significant differences associated with delay for either correct recall or incorrect recall (see Table 2). As reported earlier, when the number-of-interviews variable was examined, the main effect of age was not significant for correct responses, but was significant for incorrect statements.

Person, Room, and Action Questions

Specific questions. When the proportion of correct answers to the specific questions were considered, a MANOVA failed to reveal significant effects involving the delay factor, although a significant main effect of age duplicated

the age effect reported earlier for specific questions (see Table 3). Analyses of the proportion of omission errors also failed to reflect significant effects of delay. However, delay influenced the proportion of commission errors made. Among children who had their first interview after a 2-week delay, older children made fewer commission errors than younger children. In contrast, there were no significant age differences for children whose first interview was at 4 weeks. Main effects of delay and age were nonsignificant.

Misleading questions. When correct responses to the misleading questions were analyzed, there were no significant effects associated with delay or age (see Table 4). However, a significant main effect of delay for omission errors was subsumed under a significant Age $\times$ Delay interaction. For children whose first interview was at 4 weeks, younger children made more omission errors than older children. Furthermore, younger children first interviewed at 4 weeks made more omission errors than younger children first interviewed at 2 weeks. The proportion of omission errors made by older children at the 2-week interview did not significantly differ from that made by younger children at the 2-week interview, however. Delay did not significantly affect the proportion of commission errors made to the misleading questions. For commission errors, only the main effect of age was significant.

Abuse Questions

In response to the abuse questions, there were no significant effects associated with delay for the proportion of correct answers, omission errors, or commission errors. However, main effects of age were again significant for correct answers, omission errors, and commission errors.

Photo Identification

Finally, the 2-week photo identification performance of the children who had two interviews was compared with the 4-week performance of the children who had only one interview to determine whether there were significant differences due to delay of the first interview. A 2 (Age) $\times$ 2 (Delay) analysis revealed no significant effects involving delay or age. A 2 (Age) $\times$ 2 (Practice) analysis revealed only a significant main effect of practice in the proportion of false identifications made (no practice, $M = .75$; practice, $M = .44$), $F(1, 66) = 9.21$, $p < .01$.

In summary, although delay did not have a pervasive effect on the children's performance, it did interact with age to affect the frequency and type of errors made. Although older children made fewer commission errors than younger children to the specific questions when the first interview occurred after 2 weeks, the age advantage disappeared when the first interview occurred after 4 weeks. In contrast, age differences in omission errors to the misleading

questions were more likely to occur at 4 weeks than at 2 weeks. Thus, over time, younger children were more likely to omit information in response to the misleading questions but were not any more likely to make commission errors to these questions. On the photo identification task, delay did not exacerbate the children's initially poor performance, whereas practice led to improved performance.

Stress and Accuracy

We were also interested in the effects of stress on the children's reports. Immediately after the children received their inoculations, parents rated the children's inoculation stress. We were thus able to examine how the parents' ratings related to the accuracy of the children's reports.

When the number-of-interviews and delay variables were considered, relatively few children in each experimental condition were rated as highly stressed or unstressed. It was therefore necessary to combine into one group children who obtained a stress rating of 4, 5 and 6, and also to combine into one group children who received a rating of 1 or 2. A sufficiently high number of children received a stress rating of 3 for this category to be maintained on its own.

A series of 2 (Number of Interviews) $\times$ 3 (Stress Level) ANOVAs and a series of 2 (Delay) $\times$ 3 (Stress Level) ANOVAs were conducted on the main dependent measures (correct and incorrect free recall; proportion of correct responses, omission errors, and commission errors to the specific, misleading, and abuse questions; and false identifications of the photo identification task). Only one significant stress effect emerged and it was not immediately interpretable. For proportion correct answers to the specific questions, a significant Delay $\times$ Stress Level interaction, $F(2, 61) = 4.50$, $p < .025$, resulted from children who received a stress rating of 3 being more accurate if first interviewed at 2 weeks ($M = .76$) than at 4 weeks ($M = .56$), $F(1, 13) = 15.22$, $p < .01$ (simple effects). Given the number of tests conducted, it is possible that the one significant difference was due to chance factors.

DISCUSSION

Consistent with previous research, age differences in children's ability to report a real-life event were obtained. In fact, the most pervasive finding of this study concerned age differences. Older children were more accurate than younger children on most of the eyewitness testimony measures. Older children's accuracy did not always translate, however, into significant age differences in the types of errors made.

Despite age differences, a variety of factors led to improvement in children's reports. In general, interviewing children multiple times, providing them with

"reinforcement" and training them to better understand the identification task improved their reports. Thus, it is possible to optimize adult–child negotiation in reconstructing a child's past, resulting in a more accurate construction of reality.

Children who experienced two interviews were more accurate in answering specific questions than children who experienced a single, delayed interview. Moreover, repeated interviewing was not associated with increased suggestibility, as indexed by answers to misleading questions. Children who were interviewed twice were less suggestible at the second interview than at the first interview.

Positive effects of repeated interviewing were also evident in regard to abuse questions. Even though we suggested to the children that actions related to abuse occurred, the children were, if anything, less likely to say that something happened that did not when questioned a second time, 2 weeks later. Thus, on average, multiple interviews did not result in suggestions of abuse intruding into children's subsequent answers to our questions.

It is likely that the children in the repeated interview condition, who experienced the first interview after only 2 weeks, had an advantage because the first interview was, in effect, a chance to rehearse the information while it was still relatively fresh in memory. Thus on their second interview at the 4-week delay they were more accurate and more resistant to suggestion. In contrast, the children in the single interview condition waited 4 weeks before they were questioned. Their memory would have been more likely to fade after this longer delay.

Nevertheless, the children's memory showed surprisingly little fading from 2 to 4 weeks. Perhaps most of the forgetting that was to take place in the course of 4 weeks occurred during the first 2 weeks after the event. Although delay did not affect answers to any of the other types of questions, delay did interact with age to affect the proportion of errors made to the action questions. This may have reflected the relatively high accuracy children initially displayed in response to specific and even misleading questions about actions that did or did not occur. With such high initial accuracy, more forgetting was possible.

The positive effects of repeated interviewing despite the null effects of delay might at first seem contradictory. If the children did not significantly forget information, why should repeated testing bolster the accuracy of their reports? As Brainerd et al. (in press) have argued, testing in itself can facilitate memory. It provides an impetus to recall and rehearse the event and cues to reinstate memory.

The effects of reinforcement were of special interest. It was hypothesized that if intimidation can make child witnesses more suggestible, the converse would be true: Young children's resistance to suggestion would be strengthened with less intimidating interview techniques. Although significant effects of reinforcement were sporadic, when they did occur, they were almost always

associated with improved performance. Specifically, children who received reinforcement from interviewers, compared to those who did not, made fewer inaccurate statements in free recall. In addition, children interviewed in the reinforcement, compared to the no-reinforcement, condition made fewer commission errors in response to misleading questions after a 4-week delay. Perhaps of most importance, the use of reinforcement eliminated age differences in falsely saying something happened in response to the questions about abuse. Despite the fact that younger, compared to older, children made significantly more errors of commission to the abuse questions in the no-reinforcement condition, the age difference disappeared in the reinforcement condition.

These findings suggest that, especially when working with young children, it is important to maintain a warm, supportive interviewing style so that children feel comfortable enough to counter an adult's false suggestion and recount events accurately. Although clinicians and others have incorporated reinforcement into their interviews for many years, as we have mentioned, there have been recent attacks on interviewers for providing children with supportive comments, cookies and juice, and the like during child-abuse investigations (Underwager & Wakefield, 1990).

It is not possible from our study to determine if a specific component of our reinforcement technique was responsible for the effects we uncovered. For example, further research is needed to determine if compliments alone were responsible for the beneficial reinforcement effects we found. It could even be argued that an increase in blood sugar from the cookies and juice led to the improvement we observed. If true, one might expect a quite general increase in accuracy as a function of reinforcement. However, reinforcement did not affect the children's ability to provide correct information in free recall, to answer the specific questions accurately, or to inhibit guessing on the photo identification task. Instead, the positive effects of reinforcement were largely confined to heightening resistance to suggestion, as would be expected if this form of social support lowered the intimidation value of the interviewer.

Even without reinforcement, the children were generally accurate in answering the abuse questions, producing relatively few commission errors. However, some of the errors that were made could lead to false suspicion or perhaps even false accusations of abuse. For example, a small number of children falsely agreed that the nurse did not have any clothes on, that the child himself or herself had undressed, or that their bottoms had been touched. But, these errors were typically restricted to nods of the head or a simple *yes* answer, as opposed to elaborated responses or detailed accusations. Sometimes such errors were repeated on the second interview (e.g., a child who at both interviews responded *yes* to the question, "She didn't have any clothes on, did she?"; see also Tucker et al., 1990). When false detail was provided, it often resulted from the child recounting a previous doctor's visit.

Examples of the children's responses are illuminating. A typical 5 year old's responses are revealed in the following interview sequence:

Q. The person who gave you the shot made you take your clothes off, didn't she? *A. No.* (Correct response)
Q. Did she put anything in your mouth? *A. No.* (Omission error)
Q. Did she make you swallow some medicine? *A. Yeah.* (Correct response)
Q. She asked you to tickle her, didn't she? *A. No.* (Correct response)
Q. How many times did she hit you? *A. Didn't hit me though.* (Correct response)
Q. When she gave you the shot, did she tell you to sit still? *A. Yeah.* (Correct response)
Q. Did she take a picture of you? *A. No. She didn't have a picture-taker in there.* (Not scored)
Q. Did she talk to your mom or dad? *A. Yeah.* (Correct response)
Q. Did she touch you anywhere other than your leg? *A. Nope.* (Omission error)
Q. Did she touch your bottom? *A. No. She just had to give me a shot in my arm up here.* (Correct response)

This 5 year old was interviewed in the no-reinforcement condition. His responses can be compared with those of a 3-year-old boy also in the no-reinforcement condition:

Q. The person who gave you the shot made you take your clothes off, didn't she? *A. No.* (Correct response)
Q. Did she put anything in your mouth? *A. No.* (Omission error)
Q. Did she make you swallow some medicine? *A. Yes.* (Correct response)
Q. She asked you to tickle her, didn't she? *A. Yes.* (Commission error)
Q. How many times did she hit you? *A. Five.* (Commission error)
Q. When she gave you the shot, did she tell you to sit still? *A. Yes.* (Correct response)
Q. Did she take a picture of you? *A. Yes.* (Not scored)
Q. Did she talk to your mom or dad? *A. No.* (Omission error)
Q. Did she touch you anywhere other than your leg? *A. No.* (Omission error)
Q. Did she touch your bottom? *A. Yes.* (Commission error)

One can compare this child's report to the following from a 4-year-old boy in the reinforcement condition:

Q. The person who gave you the shot made you take your clothes off, didn't she? *A. No.* (Correct response)
Q. Did she put anything in your mouth? *A. Yes.* (Correct response)
Q. Did she make you swallow some medicine? *A. No.* (Omission error)

Q. She asked you to tickle her, didn't she? *A. No.* (Correct response)
Q. How many times did she hit you? *A. She didn't hit me.* (Correct response)
Q. When she gave you the shot, did she tell you to sit still? *A. Yes. I just had a shot with the needle and I didn't cry.* (Correct response)
Q. Did she take a picture of you? *A. No. She didn't bring a camera.* (Not scored)
Q. Did she talk to your mom or dad? *A. No.* (Omission error)
Q. Did she touch you anywhere other than your leg? *A. No.* (Omission error)
Q. Did she touch your bottom? *A. No.* (Correct response)

As can be seen, the children's responses to specific and misleading questions, including the abuse questions, tended mostly to be simple, one-word responses, regardless of age or accuracy.

Interestingly, although the 5 year old whose interview was just presented was generally quite accurate, he made an omission error when asked about something being put in his mouth. Nevertheless, when asked a more specific question—whether he swallowed anything—the child's answer was correct. The 4 year old in the reinforcement condition made the opposite error. Thus, the phrasing of a question and how children interpreted it, which was at times quite unpredictable, seemed to affect their accuracy. Such inconsistencies may have an important effect on children's perceived credibility (e.g., Leippe & Romanczyk, 1987) and adults' ability to reach the truth about what happened.

Perhaps the most inaccurate reports came from the few children who recalled a former visit to a doctor's office rather than the visit of interest. These children's responses often contained more detail than the typical inaccurate nods or single words. Thus, this conflation of memories often resulted in recounting inaccurate detail. For example, one 5-year-old boy, who recalled a man giving him a shot during a visit to a different doctor's office, was frequently incorrect and gave many pat answers that followed an obviously incorrect pattern, as exemplified here:

Q. Did the person who gave you the shot take your temperature? *A. Yes.* (Commission error)
Q. That person didn't touch you, did she? *A. No.* (Omission error)
Q. Did she put something wet on your arm? *A. Yes, alcohol* (Correct response)
Q. Did she look inside your ears? *A. Yes, it hurt.* (Commission error)
Q. The nurse who gave you the shot didn't talk to you, did she? *A. He, about 4 or 6 times.* (Partially correct response)
Q. When she gave you the shot, were you standing, sitting, or lying down? *A. Lying down—about 6 years.* (Correct response regarding lying down)
Q. The nurse kissed you on the mouth, didn't she? *A. Yes, about 5 times.* (Commission error)
Q. Did she listen to your heartbeat? *A. Yes, about 5 times.* (Commission error)
Q. The person who gave you the shot made you take your clothes off, didn't she? *A. No, only my shirt.* (Correct response)

This example points to the importance of adults making sure that they and the child are talking about the same event and that the child is attentive. Otherwise, substantial error might result.

Despite these and other errors, it is important to keep in mind that most of the children's responses to abuse questions were accurate, and that, although individual differences existed, the commission error rate was relatively low overall.

In regard to the photo identification task, the obtained effects of practice suggest that children need help in understanding that a "culprit" may not be pictured in a photo lineup. Significant improvement in children's understanding of the task occurred with only three practice trials. The beneficial effects of practice were most evident with the older children, however. It is possible that the younger children needed more practice for their performance to improve significantly.

We predicted that high levels of stress would lead to better recall and greater resistance to suggestion. Instead, when correlated and potentially confounding factors were controlled, there were virtually no significant effects of stress on the children's reports. There are several possible explanations for the lack of significant findings. Children's willingness to recall a stressful event may be one important factor. At least a few of the children rated as highly stressed seemed unwilling to talk about the visit to the clinic. It is also possible that the attained stress levels reached were not sufficiently high to produce positive or negative effects of stress on memory. This problem was exacerbated by the fact that, within each experimental group, few children received a rating of 6. Recent physiological data suggest that positive effects of stress depend on a surge of adrenalin (e.g., Gold, 1987). Given nurses' training to try to alleviate children's distress, perhaps too few of our children reached sufficiently high stress levels. Alternatively, the parents may not have been attuned to the level of stress their children were experiencing. Future studies would profit from inclusion of additional measures of stress. Finally, it is possible that the type of stress we investigated simply does not influence children's memory.

Several caveats should be issued regarding the generalizability of our findings. Concerning repeated interviewing, children in our study did not experience sexual or physical abuse and were not enmeshed in the legal system. They thus had little reason to be embarrassed by our questioning or to harbor the same degree of distress about the event and its ramifications as might afflict a child-abuse victim. Repeatedly interviewing actual child victims may cause them further distress (Tedesco & Schnell, 1987) and lead them to refuse to talk about it at all, regardless of their memory of the event. Thus repeated interviewing may be contraindicated in an actual case.

Concerning our reinforcement findings, we did not selectively reinforce children's responses to, for example, the abuse questions. We thus do not know

whether we would have been able to increase false reports of abuse if we had selectively reinforced errors.

In considering our findings regarding the effects of practice on children's photo identifications, it should be kept in mind that the implementation of these procedures in actual cases should be based on additional research to ensure that the findings are replicable. Moreover, it would be important to demonstrate that practice such as we used does not result in poorer performance on lineups where the target is present.

Finally, regarding the entire study, we did not take into account a number of factors that may influence children in actual child-abuse cases. For example, we did not pose as police or social workers who would have the power to arrest a suspect or take a child from the home. The children presumably were not motivated by revenge or the desire to protect a loved one.

Nevertheless, our findings may have important implications for the investigation of child abuse. The results imply that it may be relatively easy to develop techniques to improve children's reports, and that, rather than focusing exclusively on children's inaccuracies, we should direct more effort toward optimizing children's performance.

ACKNOWLEDGMENTS

This research was funded in part by Grant #90CA1264 to Gail S. Goodman from the National Center on Child Abuse and Neglect.

We thank Marilyn Shahan, Head Nurse, Dr. Franklin Judson, Director, and nurses Susan Minarik, Ann Schempf, Mayme Tagawa, and Sherry Wiggins, all of the Immunization Clinic of the Colorado State Department of Health and Hospitals, for their support and involvement. Christine Aman, Patricia England, Linda Port, Sherry Thomas, and Mary Beth Wilson provided valuable research assistance.

REFERENCES

Achenbach, T. M., & Edelbrock, C. (1983). *Manual for the Child Behavior Checklist and the Revised Child Behavior Profile.* Burlington: University of Vermont.

Baddeley, A. D. (1982). *Your memory: A user's guide.* New York: Macmillan.

Baker-Ward, L., Hess, T. M., & Flannagan, D. A. (1990). The effects of involvement on children's memory for events. *Cognitive Development, 5,* 55–70.

Bohannon, J. N. (1988). Flashbulb memories for the space shuttle disaster: A tale of two theories. *Cognition, 29,* 179–196.

Brainerd, C., & Ornstein, P. A. (in press). Children's memory for witnessed events: The developmental backdrop. In J. Doris (Ed.), *Children's suggestibility (with special reference to the child witness).* Washington DC: American Psychological Association.

Brainerd, C., & Reyna, V. (1988). Memory loci of suggestibility development: Comment on Ceci, Ross, & Toglia. *Journal of Experimental Psychology: General, 117,* 197–200.

Brainerd, C., Reyna, V. F., Howe, M. L., & Kingma, J. (1990). Development of forgetting and reminiscence. *Monographs of the Society for Research in Child Development, 55* (Serial No. 222).

Brigham, J. C., Maass, A., Martinez, D., & Whittenberger, Q. (1983). The effect of arousal on facial recognition. *Basic and Applied Social Psychology, 4,* 279–293.

Brigham, J. C., Van Verst, M., & Bothwell, R. K. (1986). Accuracy of children's eyewitness identification in a field setting. *Basic and Applied Social Psychology, 7,* 295–306.

Brown, R., & Kulik, J. (1977). Flashbulb memories. *Cognition, 5,* 73–99.

Case, R. (1985). *Intellectual development.* New York: Academic.

Ceci, S. J., Ross, D. F., & Toglia, M. P. (1987). Age differences in suggestibility: Narrowing the uncertainties. In S. J. Ceci, M. P. Toglia, & D. F. Ross (Eds.), *Children's eyewitness memory* (pp. 79–91). New York: Springer-Verlag.

Christianson, S., & Loftus, E. F. (1987). Memory for traumatic events. *Applied Cognitive Psychology, 1,* 225–239.

Davies, G., Tarrant, A., & Flin, R. (1989). Close encounters of the witness kind: Children's memory for a simulated health inspection. *British Journal of Psychology, 80,* 415–429.

Deffenbacher, K. (1983). The influence of arousal on reliability of testimony. In B. R. Clifford & S. Lloyd-Bostock (Eds.), *Evaluating eyewitness evidence* (pp. 235–251). Chichester, England: Wiley.

Dent, H. R., & Stephenson, G. M. (1979). An experimental study of the effectiveness of different techniques of questioning child witnesses. *British Journal of Social and Clinical Psychology, 18,* 41–51.

Easterbrook, J. A. (1959). The effect of emotion on the utilization and organization of behavior. *Psychological Review, 66,* 183–201.

Erdelyi, M. (1985). *Psychoanalysis: Freud's cognitive psychology.* New York: Freeman.

Fischer, K. W. (1980). A theory of cognitive development: The control and construction of hierarchies of skills. *Psychological Review, 87,* 477–531.

Fischer, K. W., & Bullock, D. (1984). Cognitive development in school-age children: Conclusions and new directions. In A. Collins (Ed.), *Development during middle childhood: The years from six to twelve* (pp. 70–146). Washington, DC: National Academy Press.

Fivush, R., & Hammond, N. R. (1989). Time and again: Effects of repetition and retention interval on 2-year-olds' event recall. *Journal of Experimental Child Psychology, 47,* 259–273.

Flin, R. H. (in press). A grand memory for forgetting. In J. Doris (Ed.), *Children's suggestibility (with special reference to the child witness).* Washington, DC: American Psychological Association.

Freud, S. (1975). *Three essays on the theory of sexuality.* (J. Strachey, Trans). New York: Basic Books. (Original work published 1905)

Gold, P. E. (1987). Sweet memories. *American Scientist, 75,* 151–155.

Goodman, G. S. (in press). Stress and accuracy in research on children's testimony: Commentary on Peters. In J. Doris (Ed.), *The suggestibility of children's memory (with special reference to the child witness).* Washington, DC: American Psychological Association.

Goodman, G. S., & Aman, C. (in press). Children's use of anatomically detailed dolls to recount an event. *Child Development.*

Goodman, G. S., Aman, C., & Hirschman, J. (1987). Child sexual and physical abuse: Children's testimony. In S. J. Ceci, M. P. Toglia, & D. F. Ross (Eds.), *Children's eyewitness memory* (pp. 1–23). New York: Springer-Verlag.

Goodman, G. S., & Hahn, A. (1987). Evaluating eyewitness testimony. In I. Weiner & A. Hess (Eds.), *Handbook of forensic psychology* (pp. 258–292). New York: Wiley.

Goodman, G. S., Hirschman, J., Hepps, D., & Rudy, L. (in press). Children's memory for stressful events. *Merrill–Palmer Quarterly.*

Goodman, G. S., & Reed, R. S. (1986). Age differences in eyewitness testimony. *Law and Human Behavior, 10,* 317–332.

Goodman, G. S., Rudy, L., Bottoms, B. L., & Aman, C. (1990). Children's concerns and memory: Issues of ecological validity in children's testimony. In R. Fivush & J. Hudson (Eds.), *Knowing and remembering in children* (pp. 249–284). New York: Cambridge University Press.

Harsch, N., & Neisser, U. (1989, November). *Substantial and irreversible errors in flashbulb memories of the Challenger explosion.* Paper presented at the meeting of the Psychonomic Society, Atlanta, GA.

Heuer, F., & Reisberg, D. (1990). Vivid memories of emotional events: The accuracy of remembered minutiae. *Memory and Cognition, 18,* 496–506.

Idaho v. Wright, 116 Idaho 382, 775 P. 2d 1224 (1989).

Kety, S. D. (1970). The biogenic amines in the central nervous system: Their possible role in arousal, emotion, and learning. In F. Schmitt (Ed.), *The neurosciences: Second study program* (pp. 324–336). New York: Rockefeller University Press.

King, M. A., & Yuille, J. C. (1987). Suggestibility and the child witness. In S. J. Ceci, M. P. Toglia, & D. Ross (Eds.), *Children's eyewitness memory* (pp. 24–35). New York: Springer-Verlag.

Leippe, M. R., & Romanczyk, A. (1987). Children on the witness stand: A communication/persuasion analysis of jurors' reactions to child witness. In S. J. Ceci, M. P. Toglia, & D. F. Ross (Eds.), *Children's eyewitness memory* (pp. 155–177). New York: Springer-Verlag.

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

Loftus, E. F., & Burns, T.E. (1982). Mental shock can produce retrograde amnesia. *Memory & Cognition, 10,* 318–323.

Marin, B. V., Holmes, D. L., Guth, M., & Kovac, P. (1979). The potential of children as eyewitnesses. *Law and Human Behavior, 3,* 295–306.

McCarthy, D. (1972). *Manual for the McCarthy Scales of children's abilities.* New York: Psychological Corporation.

McCloskey, M., Wible, C., & Cohen, N. (1988). Is there a special flashbulb-memory mechanism? *Journal of Experimental Psychology: General, 117,* 171–181.

McGaugh, J. L. (1989). Modulation of memory storage processes. In P. R. Solomon, G. R. Goethals, C. M. Kelley, & B. R. Stephens (Eds.), *Memory: Interdisciplinary approaches* (pp. 33–64). New York: Springer-Verlag.

Melamud, B. G. (1976). Psychological preparations for hospitalizations. In S. Rachman (Ed.), *Contributions to medical psychology* (pp. 43–74). Oxford: Pergamon.

Moston, S. (1987, September). *The effects of the provisions of social support in child interviews.* Paper presented at the meeting of the British Psychological Association, York, England.

Myers, J. (1987). *Child witness law and practice.* New York: Wiley.

Neisser, U. (1982). Snapshots or benchmarks? In U. Neisser (Ed.), *Memory observed: Remembering in natural contexts* (pp. 43–48). San Francisco: Freeman.

Oschner, J. E., & Zaragoza, M. S. (1988, March). *The accuracy and suggestibility of children's memory for neutral and criminal eyewitness events.* Paper presented at the meeting of the American Psychology and Law Association, Miami.

Parker, J. F., & Carranza, L. E. (1989). Eyewitness testimony of children in target-present and target-absent lineups. *Law and Human Behavior, 13,* 133–150.

Peters, D. P. (1987). The impact of naturally occurring stress on children's memory. In S. J. Ceci, M. P. Toglia, & D. F. Ross (Eds.), *Children's eyewitness memory* (pp. 122–141). New York: Springer-Verlag.

Peters, D. P. (1989, April). Stress and arousal effects on the child eyewitness. In C. Brainerd (Chair), *Children's ability to remember witnessed events: Theoretical and applied perspectives.* Symposium conducted at the meeting of the Society for Research in Child Development, Kansas City, MO.

Piaget, J. (1952). *The origins of intelligence in children* (M. Cook, Trans.). New York: International Universities Press. (Original work published 1936)

Pillemer, D. (1984). Flashbulb memories of the assassination attempt on President Reagan. *Cognition, 16,* 63–80.

Price, D. W. W., & Goodman, G. S. (1990). Visiting the wizard: Children's memory of a recurring event. *Child Development, 61,* 664–680.

Rudy, L., & Goodman, G. S. (in press). The effects of participation on children's reports: Implications for children's testimony. *Developmental Psychology.*

Steward, M. S. (1989). *The development of a model interview for young child victims of sexual abuse: Comparing the effectiveness of anatomical dolls, drawings and video graphics* (Grant No. 90CA1332). Washington, DC: National Center on Child Abuse and Neglect

Steward, M. S., & Steward, D. S. (1981). Children's conceptions of medical procedures. In R. Bibace & M. E. Walsh (Eds.), *Children's conceptions of health, illness, and bodily functions: New directions for child development* (Vol. 14, pp. 67–84). San Francisco: Jossey-Bass.

Tedesco, J. F., & Schnell, S. V. (1987). Children's reactions to sex abuse investigation and litigation. *Child Abuse and Neglect, 11,* 267–272.

Tucker, A., Mertin, P., & Luszcz, M. (1990). The effect of a repeated interview on young children's eyewitness memory. *Australian and New Zealand Journal of Criminology, 23,* 117–124.

Underwager, R., & Wakefield, H. (1990). *The real world of child interrogations.* Springfield, IL: Thomas.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes* (M. Lopez-Morillas & M. Cole, Trans.). Cambridge, MA: Harvard University Press. (Original work published 1934)

Warren, A., Hulse-Trotter, K., & Tubbs, E. C. (in press). Inducing resistance to suggestibility in children. *Law and Human Behavior.*

Warren-Leubecker, A., Bradley, C., & Hinton, I. (1988, March). *Scripts and the development of flashbulb memories.* Paper presented at the Conference on Human Development, Charleston, SC.

Wolfe, D. A., & Mosk, M. D. (1983). Behavioral comparisons of children from abusive and distressed families. *Journal of Consulting and Clinical Psychology, 51,* 702–708.

Yerkes, R. M., & Dodson, J. D. (1908). The relation of strength of stimulus to rapidity of habit formation. *Journal of Comparative Neurology and Psychology, 18,* 459–482.

Yuille, J. C., & Cutshall, J. L. (1986). A case study of eyewitness memory of a crime. *Journal of Applied Psychology, 71,* 291–301.

Zaragoza, M. S. (1987). Memory, suggestibility, and eyewitness testimony in children and adults. In S. J. Ceci, M. P. Toglia, & D. F. Ross (Eds.), *Children's eyewitness memory* (pp. 53–78). New York: Springer-Verlag.

Zaragoza, M. S., & Wilson, D. (1989, April). Suggestibility in the child witness. In G. S. Goodman (Chair), *Do children provide accurate eyewitness reports? Social policy and research implications.* Symposium presented at the meeting of the Society for Research in Child Development, Kansas City, MO.