

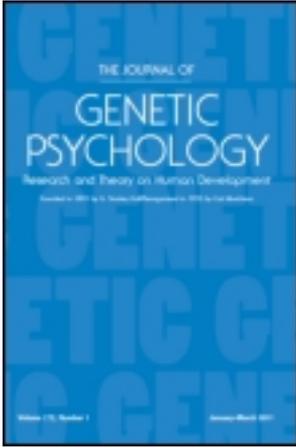
This article was downloaded by: [Marc A. Lindberg]

On: 13 February 2012, At: 03:03

Publisher: Routledge

Informa Ltd Registered in England and Wales Registered Number: 1072954

Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



## The Journal of Genetic Psychology: Research and Theory on Human Development

Publication details, including instructions for authors and  
subscription information:

<http://www.tandfonline.com/loi/vgnt20>

### Comparisons of Three Different Investigative Interview Techniques With Young Children

Marc A. Lindberg <sup>a</sup>, Mary Tantalo Chapman <sup>a</sup>, David  
Samsock <sup>a</sup>, Stuart W. Thomas <sup>a</sup> & Anders W. Lindberg <sup>a</sup>

<sup>a</sup> Department of Psychology, Marshall University

Available online: 30 Mar 2010

To cite this article: Marc A. Lindberg, Mary Tantalo Chapman, David Samsock, Stuart W. Thomas & Anders W. Lindberg (2003): Comparisons of Three Different Investigative Interview Techniques With Young Children, *The Journal of Genetic Psychology: Research and Theory on Human Development*, 164:1, 5-28

To link to this article: <http://dx.doi.org/10.1080/00221320309597500>

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: <http://www.tandfonline.com/page/terms-and-conditions>

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae, and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand, or costs or damages whatsoever or howsoever

caused arising directly or indirectly in connection with or arising out of the use of this material.

## Comparisons of Three Different Investigative Interview Techniques With Young Children

MARC A. LINDBERG  
MARY TANTALO CHAPMAN  
DAVID SAMSOCK  
STUART W. THOMAS  
ANDERS W. LINDBERG  
*Department of Psychology  
Marshall University*

---

**ABSTRACT.** After viewing a film of a mother hitting her son, a film not seen by the college student interviewers, children were misinformed about a detail (via exposure to a misleading question) as well as explicitly coached to disclose 3 false details. The children were then interviewed by interviewers who had previously learned 1 of 3 different interviewing procedures: the Yuille Step-Wise Interview developed by J. C. Yuille, R. Hunter, R. Joffe, & J. Zaparniuk (1993); a doll play interview developed by Action for Child Protection Inc. (1994); or the Modified Structured Interview developed for this study. The Modified Structured Interview yielded more “where” information and was better at detecting if coaching had occurred. However, the interviewers were not very good at discriminating suggested versus coached versus correct witnessed information. The authors found that the deeper one digs for memories, the more one uncovers incorrect versus correct items. They concluded that although the Modified Structured Interview was superior to the techniques currently in use, cautions are necessary.

**Key words:** eyewitness testimony, forensic interviewing of children, suggestibility

---

ALTHOUGH THERE IS A WEALTH OF INFORMATION about children’s memory and eyewitness testimony, we are just beginning to understand the relative effectiveness of procedures currently being used to question young children in real cases of physical or sexual abuse (Poole & Lamb, 1998). Can we develop retrieval devices that will enable us to determine the individual events that a child has experienced? To use a Freudian metaphor, can we develop a Rosetta stone that will enable us to unlock the hieroglyphics of a child’s past, or is that a faulty metaphor for how memories are stored? Our purpose in the present study was to shed light on answers to four questions. First, what interviewing technique for children gives us the most correct versus incorrect, coached, and suggested infor-

mation? Second, how is information that is provided by child witnesses received and interpreted by interviewers? Third, what implications do these relations have in terms of the practice of training new interviewers? Fourth, what do these results have to say about theories of memory in general?

One of the most comprehensive studies to date comparing different interview procedures for children is a study by Steward and Steward (1996). They created four experimental interviews that involved simply talking, talking plus dolls, talking plus line drawings, or talking plus computer graphics. The children went through a medical examination and then were interviewed with one of the four techniques. Steward and Steward (1996) did not find significant differences among the techniques in terms of the children's accuracy or consistency. The only differences found were for the "completeness" data, with marginal advantages for the enhanced interviews. However, the enhanced interviews also were associated with more false allegations of anal touches with dolls as prompts, and more genital touches when drawings were used.

Although the scope of Steward and Steward's (1996) investigation was impressive, it contains the following methodological and conceptual problems. First, the experimenters had a general idea of what the children went through, unlike interviewers in the "real world." This could have inflated recall through experimenter bias, whereby the interviewers could have used cues and hints that are not typically known in real investigations (Bruck & Ceci, 1996; Ceci & Bruck, 1996). Second, the interviewers knew all four techniques; that knowledge could have inadvertently increased effort on a favorite. Third, the researchers did not test protocols that are typically used to investigate cases of child abuse (McGough, 1996). Finally, they gave no suggestions and did not attempt to determine if coaching occurred, something that is of crucial importance in legal settings (Ornstein, 1996). In real cases of child abuse, there is always the specter of suggestions, and one often has to attempt to sift through real versus fabricated memories or at least address these issues for litigation (Melton et al., 1995).

Several other training studies have been conducted. However, many of these studies have contained the same types of flaws or have compared their technique with some obviously inferior technique that has not been well documented, such as those techniques used by the police (Geiselman, Saywitz, & Bornstein, 1993; Sternberg, Lamb, Orbach, Esplin, & Mitchell, 2001). In other training studies, the rapport-building process and the more directive questioning were typically var-

---

*The authors thank Maggie Bruck, Deborah Poole, Alison Clark-Stewart, and Amy Warren for reviewing this article. They also thank the principals, teachers, and students at Highlawn and Miller Elementary Schools for their cooperation. Part of this project was accomplished through the granting of partial release time from the university to the first and fourth authors. For copies of the Modified Structured Interview technique with the interview sheets, contact Marc A. Lindberg, who holds the copyright.*

*Address correspondence to Marc A. Lindberg, Department of Psychology, Marshall University, Huntington, WV 25755; lindberg@marshall.edu (e-mail).*

ied simultaneously. For example, Fisher and McCauley (1995) compared the cognitive interview procedure (Geiselman et al., 1993) with what they termed a standard interview. The cognitive interview technique was found to be superior to control conditions. However, in addition to having many of the same problems as Steward and Steward's (1996) study, the procedures in Fisher and McCauley's (1995) study differed with respect to the rapport-building, free-recall, and questioning phases. Bull (1995) reasoned that we are therefore left with the problem of deciding which phase of the interrogation produced the results that Fisher and McCauley (1995) obtained. Furthermore, if we rely on accuracy rates without including items from coached or suggested information, then generalization to forensic settings in which such concerns are of paramount importance will be difficult. This possibility was emphasized by Orbach et al. (2000), who suggested that laboratory studies are essential in our search for better interview techniques.

In line with this emphasis, some researchers assessing interviewing techniques have begun to study the rapport-building and child-preparation phase. One phenomenon that has been discovered is that if the interviewers begin with practice recall attempts of information unrelated to the target event in the rapport-building part of the interview, the children will recall significantly more information in free open recall (Sternberg et al., 1997; Warren et al., 1999). The least studied and most controversial area, however, has been the analysis of different types of follow-up question strategies after the first open-recall attempt. Although there is a fair amount of consensus that one should emphasize the first open question (Sternberg et al., 2001), the follow-up questions are usually the ones that dominate most forensic interviews (Hershkowitz, Lamb, Sternberg, & Esplin, 1997; Lamb et al., 1996; Sternberg et al., 1996; Sternberg et al., 1997). This seems to be especially true of children: It is often assumed that they need more prompts to elicit relevant details because of some hypothesized retrieval deficit hypothesis (Orbach et al., 2000).

Our purpose in the present study was to test the follow-up-questioning procedures of the major interviewing techniques currently being used in forensic settings with the technique developed here. We also attempted to correct for many of the problems found in the aforementioned studies. To obtain a better comparison of the three interview techniques in terms of how they probe memory, we equated rapport-building skills and procedures for the three interview groups and introduced variations only during the directed recall phases. To avoid prior exposure effects on the interviewers when the results of one interview could inadvertently spill over into the results of the next, we ensured that the number of interviewers equaled the number of children being interviewed so that each interviewer questioned only one child. We also attempted to understand the strengths and weaknesses of children's testimony as currently elicited in realistic forensic interviews already in use rather than compare them with an inadequate technique known to be deficient or with a nonexistent technique used for experimental purposes only. Although this was a risky strategy in that many researchers find few technique differences in this more directive phase (Orbach et al., 2000;

Steward & Steward, 1996), we deemed it more important to know similarities and differences between the techniques currently in use and the one developed here. Further, we attempted to compare the different techniques' tendencies to elicit but also detect suggested and coached information. Finally, we sought to test correlations between the interviewer's perceptions of the accuracy and completeness of children's responses, and the real accuracy and completeness of those responses.

To test these more specific questions, we trained college students to conduct one of three types of forensic interviews designed for children. College students with no prior history of forensic interviewing were used because of the literature showing that although one can train experienced interviewers in new techniques, experienced interviewers often do not use them (Aldridge & Cameron, 1999; Craig, Sheibe, Kircher, Raskin, & Dodd, 1999; Davies & Wilson, 1997; Memon, Bull, & Smith, 1995; Stevenson, Leung, & Cheung, 1992; Warren et al., 1999). Although all interviewers received the same training on the rapport-building phase, they were trained in one of three different interview strategies for the questioning phase. These interview strategies were the typical child protective services (CPS) interview that uses dolls (Action for Child Protection Inc., 1994), the Yuille Step-Wise Interview (Yuille et al., 1993; J. C. Yuille, personal communication, September 1995), and the Modified Structured Interview that used who, what, when, and where diagnostic sheets developed here and similar to the approaches of Memon, Wark, Holley, Bull, and Koehnken (1997) and Saywitz (1995).

After seeing a film that contained, among other things, a mother slapping her child in the face so hard that it sent him to the floor screaming, the children were given a suggestion and also deliberate coaching by the experimenter. The experimenter offered the suggestions by orally posing the question "What was the mother saying as she was taking Marc to the kitchen to wipe the blood coming from his bloody nose?" (There was in fact no blood shown.) Coaching consisted of having the experimenter say, "I want you all to do a very special favor for me because it would help my college grades a lot if you told them (the interviewers) that the older boy was also hit with a big wooden spoon. I want you to tell them that he got hit with the big wooden spoon because he wet his pants." Shortly thereafter, the children were interviewed with one of the three techniques. The interviewers were college students who had not yet seen the film, had not learned another interview technique, and had not previously interviewed another child thereby creating possible biases in their use of the technique they were trained to perform.

In summary, in the present experiment we attempted to provide a better test of techniques currently in use with a more precise design that would test not only for quantitative differences between the techniques but also for qualitative differences in correct versus incorrect who, what, when, and where information. We also attempted to assess the techniques' relative abilities to detect suggestions and coaching. Finally, we sought to provide the important descriptive data on the strengths and weaknesses of these techniques without confounds of experimenter bias or previous knowledge.

## Method

### *Participants*

The participants were 64 children, 33 from the first grade and 31 from the second grade, enrolled in two elementary schools, with mean ages of 6.6 and 7.8 years, respectively. The 35 boys and 29 girls were White and Appalachian with heterogeneous socioeconomic backgrounds ranging from lower- to upper-middle income populations. They were from city schools where the parents ranged from professionals to those on public assistance. We did not obtain personal records from the participants and therefore could not more clearly test for any differences with regard to socioeconomic background. The interviewers were 64 college students—12 men, 40 women, and 12 with gender indication missing. They were from upper division experimental and developmental psychology university classes and participated as part of a class project. Their socioeconomic backgrounds seemed similar to those of the children in this study.

### *Interviewer Training*

Training sessions for the interviewers took place the week before the children were interviewed. Before the in-class and out-of-class training sessions, all interviewers were provided with a packet containing a brief description of the experiment and its importance, instructions on how to build rapport, and one of the three different sets of instructions on how to interview the children in the specific questioning phase. They had not been instructed about the relative merits of different interviewing strategies and were blind to the hypotheses of the study.

*Rapport training.* The rapport-building section consisted of six double-spaced pages and outlined how to welcome and talk to the child when he or she was picked up from the classroom (e.g., complimenting the child's attire, discussing the child's favorite television show). After reading the rapport-building materials, the interviewers were provided with 1-hr in-class training that included viewing a videotape demonstrating the rapport-building techniques. These videotapes contained information on how to begin by asking the children to recall a familiar event (Fisher & Geiselman, 1992; Geiselman, Saywitz, & Bornstein, 1993). The recommendations for interviewers put forth by Gordon, Schroeder, Ornstein, and Baker-Ward (1995) on memory testing and by Walker and Warren (1995) on language and communication strategies also were discussed. The film was followed by a class lecture and discussion covering the same topics. In the rapport-building videotape, all interviewers were told that the child participants viewed a film. The interviewers did not see the film and were not told anything about it except that the film did not involve sex or sexual abuse and that they were not to question the children about sex. The inter-

viewers were told that their purpose was to determine the content and details of the film through their particular interviewing technique. The videotape ended by instructing the interviewers to ask the open-ended question "I know you saw a movie earlier today that I haven't seen. I need you to tell me what you saw so I can tell my teacher." The rapport-building training sessions ended with an emphasis on rereading the packets of materials relevant to the interview style, not discussing their style with others in the class, and closely following the outlined procedures. The interviewers also were told that they would be helping to improve interview techniques, thereby potentially helping children who may have been abused. Following this, the participants were randomly assigned to a condition with the stipulation that there would be the same percentage of participants in each condition from each of the classes.

*The CPS interview.* The college interviewers in the CPS condition learned the technique typically used by social workers from the West Virginia Department of Human Services by studying a 14-page double-spaced packet of information describing the technique, which was abstracted from Action for Child Protection Inc. (1994). This packet was developed with the help of a social worker from the West Virginia Department of Human Services who had participated in several national seminars on the subject and who had trained her staff to conduct forensic interviews with children. Interviewers also viewed a 20-min videotape prepared by the senior author outlining the procedures and giving concrete examples on how to use it and by attending a 1-hr-long training session.

This CPS interview procedure was designed to be flexible and sensitive to the child's emotions. The need to be sensitive to the child and not rush the pace of the interview was stressed. The interviewer was to accomplish this by exploring any perceptions, fears, and anxieties the child might have. The videotape explained that if the interviewers were taken by surprise by what the child might say, they should say something like "and what happened next" to give them time to collect their thoughts. Similar to the other interviewing techniques tested, the CPS interview procedure taught interviewers that anxiety, repulsion, or enthusiasm exhibited in response to what the child described could lead the child to be less open (Saywitz, 1995).

The interviewers were instructed on how to use the following types of questions: (a) general questions, such as "How are you?" or "Tell me more about that"; (b) focused questions, such as "How are you getting along with your dad?"; (c) multiple-choice questions, such as "Did he hit you with his hand or a stick?"; (d) yes-no questions, such as "Did he tell you not to tell?"; and (e) leading questions, such as "He made you sleep outside, didn't he?" These questions were listed on a continuum from most reliable (general) to least reliable (leading), with a strong emphasis on using the more reliable. The interviewers were instructed to try to rely on the general and focused questions and to avoid leading questions. The videotape and written packet also explained the importance

of isolating events of abuse and encouraged interviewers to ask who, what, where, and when questions.

The CPS interview technique also included the use of dolls. The interviewer was to use dolls only to gain more information about what had been recalled and to get greater clarity and organization of what the child had witnessed after the verbal statements had been taken. Although many caseworkers use anatomically correct dolls with removable clothing, the dolls in the present experiment were not of this type. The dolls were plastic and included a man and a woman, both 6.5 cm tall; and a boy and a girl, both 4.5 cm tall. The child was asked to act out what happened in the film. The interviewer then would caution the child not to pretend or imagine when using the dolls. As the child demonstrated with the dolls, the interviewer would prompt as needed by saying "Could you show me what happened next?"

*Yuille's Step-Wise technique.* A second group of college interviewers learned the Step-Wise technique developed by Yuille, Hunter, and Joffe (1990) and Yuille et al. (1993). This technique is now the standard format used in Canada and in several other places (Poole & Lamb, 1998). Similar to the other groups, the interviewers first received an 11-page packet outlining the interview technique. This packet was abstracted from Yuille's training program for child protective workers (J. C. Yuille, personal communication, September 1995). They also viewed a training videotape prepared by Yuille (J. C. Yuille, personal communication, September 1995) describing the technique and attended a 1-hr-long training session conducted by the senior author. The Step-Wise interview was designed with three distinct goals in mind: (a) to minimize any trauma the child may experience during the interview; (b) to maximize the amount and quality of the information obtained from the child while, at the same time, minimizing any contamination of that information; and (c) to maintain the integrity of the investigative process for the agencies involved. The Step-Wise approach was so named because interviewers are instructed to begin with the most general questions and to proceed to more narrow and focused forms of questioning when required. The interviewers were told "The less prompting, the better" and were instructed to demonstrate patience.

The first step in the questioning was the free narrative, in which the interviewer asked for a full recall of the incident. Although the first general question marked the beginning for all three interview conditions, those in the Yuille technique received more instruction on techniques used in this phase and were given several examples on how to ask "What happened next?" or "You were saying that [here the interviewer was told to restate the last thing the child said] and then what happened?"

After the child had finished the free narrative for an incident, the interviewer was instructed to begin to ask open questions. The purpose of the open question was to elicit more details of the incident. For example, "Do you remember any more about the time it happened in the garage?" would be an example of an

open question with a fairly general prompt. The questions should neither be leading nor contain the answer.

The next step the training videotape and the information pack covered was the use of specific questions. The purpose of this step was to provide an opportunity to clarify and extend previous answers. The interviewers were instructed to examine what had been stated in the first two steps, to ask follow-up questions, and finally to try to get even more details about the child's earlier recall attempts. This way, they could better learn the specifics of the child's movie and the details that occurred in it.

*The Modified Structured Interview.* The third group of interviewers learned the Modified Structured Interview technique developed at Marshall University. The interviewers watched a videotape prepared by the senior author, studied a 17-page packet of information describing the technique, and attended a 1-hr-long training session. This interviewing technique contained many aspects of the Step-Wise interview procedure. Both began with very open questions and proceeded to more narrow forms of questioning only when required. As with the other two techniques, the importance of the interviewer's remaining patient when obtaining information from the child was stressed.

After the free narrative, the interviewers were instructed to begin using diagnostic question sheets rather than pictures, as has been done by Saywitz (1995), to guide more specific questioning. The purpose of the question sheets was to assist the child in recalling more who, what, when, and where details of the incident. The interviewer diagnostic question sheets consisted of five different categories. The first was the fearful-questions sheet. This sheet contained questions such as "When you watched the movie, who would you be most afraid of?" The interviewer was then to ask follow-up questions such as "What did they do that would make you afraid?" "How did they do it?" "What did they say?" "When did it happen?" and "Did anything weird or strange happen when they did it?" The second type of question was the "Who was present?" type. An example of this type of question is "Who was in the room when it happened?" This probe was followed by requests for descriptions of what each person was wearing and included requesting clothing colors. The third type of question was the how questions. These included questions about weapons and implements. The question sheet used the example, "Did they hold, touch, or hit anything?" Follow-ups for this question might include questions on size, color, or process. The fourth question sheet was used to obtain where information. Interviewers asked questions such as "Please tell me something about the room, playground, house, and so forth." These general where questions were followed with questions aimed at obtaining more details such as "What was in it?" (room, house, or elsewhere); "Where were the objects placed?" (if the child recalled objects); "Where did the (incident) take place?"; "Where were the people?"; and "Where were you?" The final question sheet attempted to detect coached information. A sample question

from this sheet included "Did anyone ask you to tell me anything about what you saw?" The interviewers were told that, when appropriate, they were to follow up with questions concerning what they were told to say and if it was something that really happened. As with the other techniques, the interviewers told the children that the response "No, I don't remember" is perfectly acceptable. If an interviewer came across any inconsistencies in the child's statements, the interviewer was instructed to return to the inconsistencies at the end of the interview and attempt to clarify them with the child. The Modified Structured Interview technique concluded in much the same way as the CPS and Yuille techniques. The interviewer thanked the child for participating, no matter what the outcome, and asked if there were any questions.

In summary, interviewers participated in one of three different training sessions. The primary differences in the techniques were not in the rapport-building or general-question phases (Bull, 1995). Rather, they differed in the more specific-questioning phase. In the CPS condition, prompting was accomplished through the use of the dolls. In the Yuille condition, prompting was accomplished through the use of progressively more focused questions and prompts on information obtained from the more general questions. Finally, in the Modified Structured Interview condition, prompting was accomplished in terms of the diagnostic question sheets.

### *Materials*

The film that was shown to the children was a 3.5-min videotape about two boys, aged 5 and 11 years old. The boys were observed coming home from school, eating snacks, engaging in conversation, playing video games, and so forth. Their mother then came home and asked the younger boy for help with a spilled bag of groceries. After many ignored requests, she came back into the room and appeared to hit him in the face, the force of the blow apparently knocking him to the floor crying.

Although it is likely children have witnessed much more violent episodes on television, it is not common for that violence to include an adult physically assaulting a child. To deal with possible negative effects of the children watching this kind of parental aggression, the interviewers debriefed the children at the conclusion of the experiment. This debriefing consisted of assuring the children that the film was made by actors and showing them one of the outtakes from the production. A film was shown in which the young boy forgot his lines and the sound effects for the simulated slap were very late, causing both characters to start laughing. It should be noted that the parental consent form specifically mentioned the fact that the mother "apparently slaps" the child. Because the children were too young to read and understand this letter and because the permission slips were distributed several weeks prior to testing the children, it is unlikely that this would have presented a confounding variable. Furthermore, when the children

came to the experimental room, they were all told that different groups would view different films and that their film was going to be seen by only their group. This study was approved by Marshall University's Institutional Board of Review.

### *Procedure*

The experiment was conducted over a time period of 2 weeks at two different elementary schools. The experiment was begun as soon as all the children arrived at school. Approximately 9–12 children who had previously been randomly assigned to each of the three groups to participate for the day were gathered together in an area away from other children (e.g., gym, auditorium, or dead-end hallway) where the film was shown.

After seeing the film, all the children were told by the experimenter that they would later be questioned about it. The experimenter then coached the children by saying,

I want you all to do a very special favor for me because it would help my college grades a lot if you told them (the interviewers) that the older boy was also hit with a big wooden spoon. I want you to tell them that he got hit with the big wooden spoon because he wet his pants. Now, let's practice recalling some things that you saw in the movie. (1) How many boys were in the film? (The correct answer was two.) (2) What did the mother say when she took Marc to the kitchen to wipe the blood coming from his bloody nose? (The correct answer was "no more minutes".)

This question about the bloody nose was a leading question because there was no blood in the filmed action. The children were then returned to their respective classrooms and were interviewed in a private space 30 to 60 min after they viewed the film.

Here it is important to distinguish between what is termed "coached" versus "suggested" information. The blood information was considered to be suggested information because of its subtle inclusion in the experimenter's questions. The information regarding the older boy being hit, the wet pants, and the wooden spoon were all considered to be coached information because of the blatant directive of the experimenter in which the children were told to say something that they did not witness.

### *Scoring*

Interviewers submitted an audiocassette as well as a typed verbatim transcript made from their audiotapes of their interviews. To analyze these transcripts, we devised a scoring system that classified the protocols into correct and incorrect statements. The incorrect statements were noted as "IC." These statements were characterized as: "who," "ICwho," "whoD," "ICwhoD," "what," "ICwhat," "when," "ICwhen," "where," and "ICwhere" information. Who information was

recorded on a transcript sheet by raters tallying how many correct actors were reported by the children and IC who referred to the number of incorrect actors listed. WhoD information was tallied by counting the number of correct adjectives describing each actor, and ICwhoD referred to the number of incorrect descriptors of the actors. What information was recorded by tallying the number of correct things and actions listed and ICwhat was tallied by counting the number of incorrect things and actions listed. When was calculated by counting the number of correct sequences listed and ICwhen referred to the number of instances things were listed out of order. Where information was tallied by counting correct room descriptions and ICwhere was tallied by counting the number of incorrect room descriptions. Coaching and other suggested information was recorded with a dichotomous yes–no or recall–no recall format. Three raters scored the data. To assess interrater reliability, all of the interviewers rated the same five randomly selected children's protocols, and their who, what, when, and where counts were correlated across the five participants. Thus, the raters represented the columns, and the child/counts on who, what, when, and where—Child 1, who, what, when, and where. Child 2, and so forth—served as the rows in the correlation matrix. The correlation between Raters 1 and 2 was .94, between 1 and 3 was .89, and between 2 and 3 was .95.

#### *College Student Perceptions Questionnaire*

A 22-item questionnaire was administered to the college-student interviewers after they had interviewed their child to measure perceptions of their child's information. The questionnaire consisted of a series of statements that the interviewers rated on a 4-point Likert-type scale: *very true*, *true*, *somewhat untrue*, and *very untrue*. Where applicable, several questions were combined to form a construct. The following represents the categories, constructs, or both used in the analyses with the number of questions representing each construct along with the alpha level: questions about whether the child gave a complete account, was honest, would be good in court, and knew the difference between suggested versus actual events. In addition, the degree to which the interview would be effective evidence was combined into one score representing the construct of rater perceptions of validity (5 questions, RVALID;  $\alpha$  of construct = .92); questions regarding interviewers' amount of interaction with same-age children, level of comfort interacting with children this age, and comfort level upon entering the interview situation represented the construct of rapport (3 questions, RAP;  $\alpha$  of construct = .60); questions asking interviewers to rate their beliefs in the existence of three coached items (3 questions, RCOACH;  $\alpha$  of construct = .66). Remaining variables from the interviewer questionnaire that were based on single items were as follows: Interviewers rated belief in the existence of blood (RBLOOD), whether the younger boy was really hit (RHIT), and how hard the children believed that they studied the materials (RSTUD).

## Results

### *Manipulation Check*

A graduate student familiar with the training procedures but blind to participant condition categorized each child's protocol in terms of which interview technique he thought had been used. He correctly identified 92% of the training conditions to which the children had been exposed. Therefore, in terms of this more pragmatic measure, one can conclude that the training of the college-student interviewers produced three clearly different types of response patterns emitted by the children. It should be parenthetically noted that we often sat outside the interview stations and listened to the interviewers. This very informal set of observations yielded very favorable impressions. The college-student interviewers seemed adept in their procedures and established good rapport with the children. The interviewers seemed to take the project seriously, and their performance matched or exceeded the performance of many of the social workers interviewing actual child victims, who were observed on videotape by the senior author. The remaining results are presented in terms of the first three questions outlined in the beginning of this article.

### *Which Technique Produced the Best Performance in Children*

The first set of analyses was designed to provide descriptive statistics for the relative outputs of the categories of dependent variables. The numbers of correct and incorrect instances recalled in both the open-ended and directed recall phases for each category of responses were analyzed by a 3 (interview condition)  $\times$  2 (child gender)  $\times$  10 (response type: correct who listed, ICwho listed, correct who descriptors, ICwho descriptors, correct what, ICwhat, correct when, ICwhen, correct where, and ICwhere) multivariate analysis of variance (MANOVA). The model was significant, Wilks's  $\lambda = .02$ ,  $F(9, 72) = 237.58$ ,  $p < .01$ .

The first follow-up MANOVA compared the 10 response types (correct who, correct what, and so forth) to test whether these different types of information were differentially produced by the children. There was a significant main effect of response type,  $F(9, 72) = 16.22$ ,  $p < .01$ . A Scheffé test revealed the following differences between the following conditions in which the prefix IC refers to incorrect responses: what = 11.89 (a), when = 8.94 (b), ICwhat = 6.34 (c), whoD = 4.37 (c,d), ICwhen = 3.36 (d), who = 2.92 (d,e), ICwhoD = 1.11 (e,f), where = 1.00 (e,f), ICwhere = .50 (f), and ICwho = .08 (f). Means that had a letter subscript in common were not significantly different.

The second MANOVA tested the key comparisons outlined in the introduction by testing for interactions of interview condition by correct versus incorrect items recalled in each of the who, whoD, what, when, and where response-type categories. This was, therefore, a 3 (interview condition)  $\times$  10 (response type) MANOVA. There was a significant interaction between response type and inter-

view condition, Wilks's  $\lambda = .45$ ,  $F(18, 98) = 2.63$ ,  $p < .01$ . Before proceeding with a discussion of follow-ups to these results, it is important to first list the assumptions behind using a MANOVA to follow up the first main MANOVA, rather than within design analysis of variance (ANOVA) tables as is more commonly found in this literature. When the children recalled the information, there were no sequential dependencies built in for emitting either correct or incorrect information or response-type information. That is, just because they could have displayed excellent recall for correct what information, this did not preclude them from recalling incorrect what information, or even correct where information. Therefore, a design assuming sequential dependency was not deemed appropriate. The MANOVA, with no assumptions of sequential dependencies, was used. The only follow-up MANOVA finding significant differences between interview conditions was for the where variable,  $F(2, 61) = 14.02$ ,  $p < .01$ . Post hoc Scheffé tests found that the Modified Structured Interview produced more correct where responses ( $M = 2.50$ ) than the Yuille Step-Wise condition ( $M = .40$ ) or CPS condition ( $M = 0.33$ ). The Modified Structured Interview was also better than the other conditions at detecting where the child had seen the movie,  $\chi^2(6, N = 60) = 18.46$ ,  $p < .01$ . No other differences were significant. (It should be noted that even if one did assume dependencies between the listed categories of recall, the same patterns and findings of statistical significance were found with within-design ANOVAs).

The third MANOVA tested possible interactions between response type and child gender. The model was significant, Wilks's  $\lambda = .70$ ,  $F(9, 49) = 2.38$ ,  $p < .03$ . Follow-up ANOVAs on the response-type-by-gender interaction showed significant gender effects for when,  $F(9, 49) = 6.32$ ,  $p < .02$ ; ICwhen,  $F(9, 49) = 5.50$ ,  $p < .02$ ; what,  $F(9, 49) = 9.15$ ,  $p < .01$ ; ICwhat,  $F(9, 49) = 6.42$ ,  $p < .01$ ; and ICwho,  $F(9, 49) = 4.68$ ,  $p = .03$ . The boys were more verbose in that they produced more correct and incorrect answers to these questions. The means for the boys on these measures were as follows: ICwhoD = 1.34, what = 13.65, ICwhat = 7.77, when = 10.37, and ICwhen = 4.31. The means for the girls on these measures were as follows: ICwhoD = .82, what = 9.75, ICwhat = 4.62, when = 7.21, ICwhen = 2.20.

The fourth MANOVA tested possible interactions between response type and age. This test did not reveal any statistically significant effects, Wilks's  $\lambda = .84$ ,  $F(9, 49) = 1.07$ ,  $p > .05$ . No other MANOVAs were significant.

In the next analyses, we attempted to determine whether the training conditions produced differences in the first free-recall question asked of all participants. That is, were interviewers in some of the training conditions more willing to begin prematurely with their specific questions than others? Because of the fewer responses emitted in this first free-recall phase, and because of the fact that some of the response categories used in the previous analyses contained zero observations, a 3 (condition)  $\times$  2 (response type: correct vs. incorrect) MANOVA was performed. A significant interaction was found, Wilks's  $\lambda = .87$ ,  $F(2, 54) = 4.03$ ,  $p = .02$ . Post hoc Scheffé tests showed higher levels for correct as compared with incorrect recall for the Modified Structured Interview and the Yuille Step-Wise

technique than for the CPS technique. For the Modified Structured Interview technique, the mean number of correct items was 13.15 and of incorrect items was 1.45. For the Yuille Step-Wise condition, the mean number of items correctly recalled during the first question was 15.90, and the mean number of items incorrectly recalled was 2.25. Finally, the mean number of correct items was 6.23 for the CPS technique, and the mean number of incorrectly recalled items was 0.82.

#### *How the Recalled Information Was Received and Interpreted by the Interviewers*

The first analysis was performed on the number of interviewers who said that their children told them that someone had told them to say some things in the interview. These "detection of coaching" data were then subjected to a 2 (grade)  $\times$  2 (correct vs. incorrect)  $\times$  3 (interview condition) chi-square analysis. The Modified Structured Interviews detected coaching in 26% of the participants as compared with only 11% in the CPS condition and 0% in the Yuille condition,  $\chi^2(6, N = 60) = 6.89, p = .03$ . The next analysis compared the conditions in terms of whether they differed in the recall of the suggestion of blood. Of those children recalling that they saw blood (37%), 83% were boys and 17% were girls, a significant finding,  $\chi^2(4, N = 63) = 11.96, p < .01$ . There were not, however, any significant differences for condition. See Table 1 for these results and others on the dichotomous data.

**TABLE 1**  
Number and Percentage Recall for Two-Way Dichotomous Variables

Variable	Condition					
	CPS		Yuille		Structured	
	No.	%	No.	%	No.	%
Where hit***	0	0	3	14	11	55
Where movie***	0	0	0	0	8	40
Wet pants young	5	25	2	9	5	26
Wet pants old	7	47	6	40	2	13
Spoon	14	67	12	52	11	58
Hit older	12	36	11	33	10	30
Hit younger	17	81	15	65	13	68
Blood	6	29	8	35	9	47
Coach*	2	11	0	0	5	26
Mom (no more min)	6	30	13	57	11	55

*Note.* CPS = Interview method typically used by child protective services, developed by Action for Child Protection Inc. Yuille = Yuille Step-Wise interview method. Structured = Modified Structured Interview technique.

\* $p < .05$ . \*\*\* $p < .001$ .

TABLE 2  
Correlations Between College Student Ratings and the Children's Performance

Variable	1	2	3	4	5	6	7	8	9	10	11
1. RVALID	—	0.18	-0.03	.19	.10	.04	.27*	.01	.29*	-.07	.18
2. RAP		—	0.33*	-.01	-.24	-.11	.11	.11	-.05	.12	.08
3. RSTUD			—	-.05	.03	.02	.14	.13	-.17	.05	-.17
4. RHIT				—	.37**	.01	.30*	.23	.37**	.43**	.12
5. RBLOOD					—	.15	.34**	.38**	.09	.79***	-.01
6. RCOACH						—	-.02	.03	.32*	.00	.46**
7. CREC							—	.36**	.11	.53**	-.02
8. ICREC								—	.09	.45**	.06
9. CHIT									—	.04	.53**
10. CBLOOD										—	-.05
11. CCOACH											—

Note. Pearson Correlation Coefficients/Prob > R under  $H_0$ ;  $p = 0$ /No. of observations. CREC = children's correct recall, including sequences, actions, descriptions and room words; ICREC = children's incorrect recall, including incorrect sequences, incorrect actions of the actors, incorrect description of room words; CHIT = children's report of the young boy being hit; CBLOOD = children's report of suggested blood; CCOACH = children's report of coached details; RBLOOD = interviewer ratings of their belief in the occurrence of blood; RCOACH = interviewer ratings of their belief in the occurrence of coached information, including the spoon and hitting of the older boy; RHIT = interviewer ratings of their belief that the young boy was hit; RVALID = interviewer ratings of whether the child gave a complete account, was honest, and would be good in court, and the degree to which the interview would be effective evidence; RAP = interviewer ratings regarding their past interactions with children of this age, whether they felt comfortable with children this age and whether they felt comfortable going into the interview; RSTUD = interviewer ratings of their preparation/study of the interview materials provided.

\* $p < .05$ . \*\* $p < .01$ . \*\*\* $p < .001$ .

The relations between the children's performance and interviewer perceptions also were analyzed in terms of correlational data. Correlations between the college-student-interviewer perception data and the children's recall data can be seen in Table 2. Dichotomous variables such as incorrect and correct were scored as 1 and 2, respectively, and recall of suggested or coached information also was coded as 1 if the children did not recall and 2 if they did recall. The degrees of freedom averaged 52.

### Discussion

Before discussing the main results, one must first address the following question: Did the different training conditions produce three different interviewing styles, and did these in turn lead the children to produce three different types of recall protocols? Because 92% of the children's protocol sheets were correctly classified by a coder who was blind to condition assignment, it would be reasonable to assume that each of the three training conditions produced three different interviewing strategies and recall styles. The 8% that were not accurately identified were ones that did not produce lengthy recall. The finding that the children receiving the Yuille and Modified Structured Interviews recalled more items during the free-recall session and the finding that children in the Modified Structured Interview group recalled more correct where information provided additional evidence that the interviewers were engaging in qualitatively different techniques and getting at different memories in different ways. Further support for the contention that the three interview techniques really involved different interview styles came from the fact that the Modified Structured Interview was better at detecting coaching by researchers. Future researchers might be interested in performing correlations between rated evaluations of interviewer behavior, training condition, children's recall, and jury perceptions. In this way, one could create more analytic path models of communication patterns from trainer, to interviewer, to child, back to interviewer, and then perhaps to perspective jurors.

The question "Which of the three interviewing methods provided the best performance in children?" was answered. The CPS interview technique with dolls was inferior to the other two techniques in the first free recall phase of the interview. The interviewers jumped in quickly with their memory aids, ending what was found to be the superior interview technique, that of just sitting and listening. The Modified Structured Interview technique, on the other hand, was better at detecting coaching and came out on top in terms of accurate where information. The Modified Structured Interview condition generated more than 6 times the number of correct room descriptors than either the CPS or Yuille Step-Wise conditions. This where information may be some of the most important taken by authorities because it allows police investigators to collect corroborating physical evidence. In psychology, we must be careful to note that it is the physical evidence that may persuade a judge or jury in a trial when allegations can often be

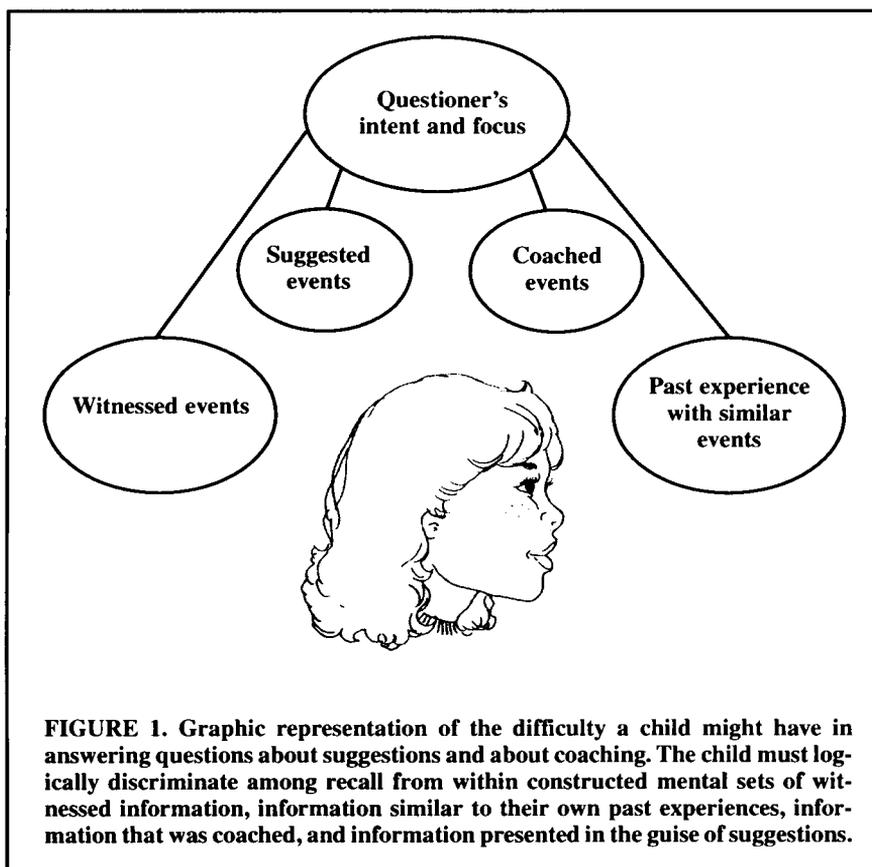
a battle between verbal statements in a she-said-he-said standoff. In our experience with actual cases, this seems to be a relatively neglected area in which many interviewers may place too much emphasis on what was recalled by the child without doing the important follow-up by the police in gathering physical evidence. By using the Modified Structured Interview, investigators may collect important eyewitness testimony about physical evidence that could help corroborate what a child has revealed.

Did the interviewing techniques differ in terms of their production of suggested or coached information? The answer was no. Although 83% of the children included suggested or coached information in their accounts (a level exceeding even the 70% recalling the central event in the film in which the younger boy got hit), there were no interview condition differences in the production of these different types of information. (See Poole & Lindsay, 1998, for an excellent review of related data on how suggested information can even enter into the open phases of questioning.)

The next question posed in this article was "How was the information provided by children received and interpreted by interviewers?" From Table 1, one can see that there were significant correlations between interviewer's ratings of validity and two other important variables, (a) number of instances correctly recalled and (b) correctly recalling the hitting. If one looked at just these data, it would appear that interviewer perceptions can be useful indexes of the validity of a child's answer, an assertion made by the Undeutsch hypothesis (Pezdek & Taylor, 2000; Raskin & Esplin, 1991; Steller, 1989). Such a conclusion would, however, be incorrect. Although the interviewers' judgments of accuracy and completeness correlated with the production of valid testimony by the children, these judgments also correlated with the production of invalid testimony, as was seen in the correlation between recall of the suggested blood and the interviewer's belief in blood,  $r = .79$ ,  $p < .001$ . Thus, the interviewers believed the children who recalled blood even though it was merely suggested and fabricated. Given the fact that even in the superior condition of the Modified Structured Interview, more than two thirds of the interviewers did not detect that coaching occurred and, if they did make such a detection, they could not determine what was coached versus suggested versus real. We should be leery of what interviewers of children believe after an interview. It must be emphasized that although there were several correlations between child performance and interviewer beliefs, the experimental safeguards in the present paradigm showed that this type of correlational data can not be used to support some version of the Undeutsch hypothesis (see Table 1). In summary, these findings, along with the results of other studies that have used experienced seasoned veterans of interviewing children in cases of abuse, alert us to the fact that interviewers are poor in their ability to determine differences between coached versus suggested versus witnessed information (Bruck, Ceci, & Hembrooke, 1998; Ceci, Loftus, Leichtman, & Bruck, 1994; Conte, Sorenson, Fogarty, & Rosa, 1991; Horner, Guyer, & Kalter, 1993; Leippe, Manion, & Romanczyk,

1993). This is troubling because interviewer beliefs are readily accepted in courts (Tobey, Goodman, Batterman-Faunce, Orcutt, & Sachsenmaier, 1995).

Why are interviewers, regardless of training, so poor at detecting suggested and coached information in this and other studies that tested very experienced interviewers? One way to explain this would be to refer to Figure 1. This figure depicts a child trying to understand and respond to the question of what was witnessed, suggested, coached, and so forth. The first thing children must do is try to figure out just what is referred to in the question. If they are able to figure out the question, then they must compare the mental sets of items from at least four different classes of information. The children then would have to independently scan each set to determine which bits of information belonged to each set. To put this in Piagetian terms, they would be required to mentally compare at least four sets of information, each with its own time tag. Children who are in Piaget's stage of concrete operations would not be expected to perform this task very



**FIGURE 1.** Graphic representation of the difficulty a child might have in answering questions about suggestions and about coaching. The child must logically discriminate among recall from within constructed mental sets of witnessed information, information similar to their own past experiences, information that was coached, and information presented in the guise of suggestions.

well. Several years ago Bransford and Franks (1971) showed that adults do not do this task very well either.

These data also have implications for issues related to questions about the nature of memory in general. A question we began with is typical of most interview-techniques research. Can we develop a Rosetta stone that will enable us to unlock the hieroglyphics of a child's past? Specifically, is it theoretically reasonable to assume a retrieval deficit hypothesis, and can we get techniques that allow us to get deeper into the truth? First, the recall data showed that as the children's recall of correct information increased, so did the amount of incorrect information. Second, it was found that the percentage of errors was more than 3 times as high on directed recall using the interview techniques as compared with free recall. Cassel and Bjorklund (1995); Dent and Stephenson (1979); Lindberg, Kiefer, and Thomas (2000); Pipe, Gee, Wilson, and Egerton (1999); Poole and Lamb (1998); Poole and Lindsay (1998); and Poole and White (1991) have made similar observations that as one moves to more directed forms of recall, the percentage of incorrect information increases. These results, when taken together, point to a potential problem for forensic investigators. Although it is desirable to get as much correct information as possible, digging deeper also increases the likelihood of more incorrect information, information that could potentially impeach the credibility of the child's testimony in court (Bull, 1995). Thus, memories may be like snow, where digging deeper only gets one into dirt, dirt that can smear a young child's mind with false memories and falsely prosecute the innocent. In conclusion, the old hope for a Rosetta stone for memory access and beliefs in retrieval deficit hypotheses as the basis of memory development seem to be based on faulty assumptions about how memory operates (Lindberg et al., 2000).

This study was also different than most in this literature in that the interviewers did not know what the children had witnessed. In classic studies conducted by Ceci, Ross, and Toglia (1987); Fisher and McCauley (1995); Rudy and Goodman (1991); and Steward and Steward (1996), interviewers generally knew what children had witnessed. Such knowledge may lead to a *Rosenthal effect*, in which children only seem to be able to recall more information when interviewers know ahead of time what happened. (See Lindberg et al., 2000, for a more complete discussion of appropriate controls and the need for analyses of multiple dependent variables.) Obviously, such work must be performed so that descriptive norms can be established documenting typical memory performance in a variety of situations, with a variety of participants, and with a variety of types of involvement. Without such descriptive studies, these and other differing forensic interview procedures will not have sound psychometric properties, properties that psychologists can more clearly use to form better expectations about children's eyewitness testimony uncontaminated by interviewer prior knowledge.

There were several shortfalls of the present research paradigm that must be noted. First, it is difficult to say if interviewers in this study represented those who work in the field. The 1 week training and preparation of these interviewers,

along with the fact that this was their first experience, might lead one to believe that they would engage in more mistakes than experienced social workers. This inexperience could have increased the error rates in these children. However, their tasks and goals (to merely find out what children had viewed in a movie), were much simpler than what typical forensic investigators face. Furthermore, it was our opinion that they were in many respects superior to experienced interviewers in the care they used to avoid suggestive questions. Finally, the data of Warren and Lane (1995) suggest that actual interviewers rarely use the same care in following new interview protocols as was exercised by the interviewers in this study and that training of experienced evaluators may have produced more conservative results. For example, Warren, Woodall, Hunt, and Perry (1996) have found that experienced interviewers tend to be more resistant to changing their already established interview behaviors.

It also should be noted that these children's memories were highly contaminated with suggestions and coaching. In examining these data alone, it might be concluded that children of this and younger ages have very poor memories in general. Such a conclusion would be very misleading. If, for example, these children had not been given the above suggestions and coaching, then their memories could have been quite good and relatively error free, mimicking those laboratory studies finding good mnemonic abilities of young children (Ornstein, Baker-Ward, Myers, Principe, & Gordon, 1995; Peterson, 1999; Pipe et al., 1999). Thus, if one wishes to use either data emphasizing excellent or poor performance of children in court, one must place this and other studies in the context of the particular case before the court.

Another criticism of the present research is that although it was more analytic than similar studies because the interview techniques were equated on rapport, it is still hard to say just what the independent variables were (cf. Ornstein, 1996). This weakness was the most obvious reason why psychological mechanisms could not be determined for each of the effects observed. However, these weaknesses also could be thought of as strengths because such an approach led to greater contextual equivalence to interview procedures typically used.

In summary, the present study afforded several conclusions. First, the dolls and memory aids did not help (DeLoache, 1995), and in terms of the first open question, they hurt recall. Second, digging deeper meant digging into more suggestions and confabulations. Third, we are probably not going to be able to tell with any certainty what was suggested, coached, or actually witnessed by the child on the basis of the child's memory alone. Fourth, interviewers should learn to remain silent and let the child take over in the first open free recall attempt, because it produces the most pristine information (Sternberg et al., 2001). Fifth, only after the open-ended questions should one use the diagnostic interview sheets of the Modified Structured Interview developed here to better understand where information and to hopefully deal with the problems of interviewers not following the training received (Poole & White, 1995; Warren, Woodall, Hunt, &

Perry, 1996). Finally, the interview is only a hypothesis-testing phase of an investigation (Poole & Lindsay, 1998). It should be used in hopes of getting more solid physical and other corroborating evidence.

## REFERENCES

- Action for Child Protection Incorporated. (1994). Training materials distributed to the Department of Health and Human Resources, 2101 Sardis Rd. N. Suite 204, Charlotte, NC 28227.
- Aldridge, J., & Cameron, S. (1999). Interviewing child witnesses: Questioning strategies and the effectiveness of training. *Applied Developmental Science*, 3, 136–147.
- Bransford, J. D., & Franks, J. J. (1971). The abstraction of linguistic ideas. *Cognitive Psychology*, 2, 331–380.
- Bruck, M., & Ceci, S. J. (1996). Commentary: Issues in the scientific validation of interviews with young children. *Monographs of the Society for Research in Child Development*, 61, Nos. 4–5, 204–214.
- Bruck, M., Ceci, S. J., & Hembrooke, H. (1998). Reliability and credibility of young children's reports. *American Psychologist*, 53, 136–151.
- Bull, R. (1995). Innovative techniques for the questioning of child witnesses, especially those who are young and those with learning disability. In M. S. Zaragoza, J. R. Graham, G. C. N. Hall, R. Hirschman, & Y. S. Ben-Porath (Eds.), *Memory and testimony in the child witness* (pp. 179–194). Thousand Oaks, CA: Sage.
- Cassel, W. S., & Bjorklund, D. F. (1995). Developmental patterns of eyewitness memory and suggestibility: An ecologically based short-term longitudinal study. *Law and Human Behavior*, 19(5), 507–532.
- Ceci, S. J., & Bruck, M. (1996). *Jeopardy in the courtroom: A scientific analysis of children's testimony*. Washington, DC: American Psychological Association.
- Ceci, S. J., Loftus, E. F., Leichtman, M. D., & Bruck, M. (1994). The possible role of source misattributions in the creation of false beliefs among preschoolers. *International Journal of Clinical and Experimental Hypnosis*, 42, 304–320.
- Ceci, S. J., Ross, D. F., & Toglia, M. P. (1987). Suggestibility of children's memory: Psychological implications. *Journal of Experimental Psychology: General*, 116(1), 38–49.
- Conte, J. R., Sorenson, E., Fogarty, L., & Rosa, J. D. (1991). Evaluating children's reports of sexual abuse: Results from a survey of professionals. *American Journal of Orthopsychiatry*, 61, 428–437.
- Craig, R. A., Sheibe, R., Kircher, J., Raskin, D. C., & Dodd, D. (1999). Effects of interviewer questions on children's statements of sexual abuse. *Applied Developmental Science*, 3, 77–85.
- Davies, G., & Wilson, C. (1997). Implementation of the memorandum: An overview. In H. Wescott & J. Jones (Eds.), *Perspectives on the memorandum: Policy, practice and research in investigative interviewing* (pp. 1–12). Aldershot, UK: Arena.
- DeLoache, J. S. (1995). The use of dolls in interviewing young children. In M. S. Zaragoza, J. R. Graham, G. C. N. Hall, R. Hirschman, & Y. S. Ben-Porath (Eds.), *Memory and testimony in the child witness* (pp. 160–178). Thousand Oaks, CA: Sage.
- Dent, H. R., & Stephenson, G. M. (1979). An experimental study of the effectiveness of differential techniques of questioning child witnesses. *British Journal of Social and Clinical Psychology*, 18, 41–51.
- Fisher, R., & Geiselman, R. E. (1992). *Memory-enhancing techniques for investigative interviewing: The cognitive interview*. Springfield, IL: Charles C. Thomas.
- Fisher, R. P., & McCauley, M. R. (1995). Improving eyewitness testimony with the cog-

- nitive interview. In M. S. Zaragoza, J. R. Graham, G. C. N. Hall, R. Hirschman, & Y. S. Ben-Porath (Eds.), *Memory and testimony in the child witness* (pp. 140–159). Thousand Oaks, CA: Sage.
- Geiselman, R. E., Saywitz, K. J., & Bornstein, G. K. (1993). Effects of cognitive questioning techniques on children's recall performance. In G. S. Goodman & B. L. Bottoms (Eds.), *Child victims, child witnesses* (pp. 71–93). New York: Guilford Press.
- Gordon, B. N., Schroeder, C. S., Ornstein, P. A., & Baker-Ward, L. E. (1995). Clinical implications of research on memory development. In T. Ney (Ed.), *True and false allegations of child sexual abuse: Assessment and management* (pp. 99–124). New York: Brunner/Mazel.
- Hershkowitz, I., Lamb, M. E., Sternberg, K. J., & Esplin, P. W. (1997). The relationships among interviewer utterance type, CBCA scores, and the richness of children's responses. *Legal and Criminological Psychology*, 2, 169–176.
- Horner, T. M., Guyer, M. J., & Kalter, N. M. (1993). Clinical expertise and the assessment of child sexual abuse. *Journal of the American Academy of Child and Adolescent Psychiatry*, 32, 925–931.
- Lamb, M. E., Hershkowitz, I., Sternberg, K. J., Esplin, P. W., Hovav, M., Manor, T., et al. (1996). Effects of investigative utterance types on Israeli children's responses. *International Journal of Behavioral Development*, 19, 627–637.
- Leippe, M. R., Manion, A. P., & Romanczyk, A. (1993). Discernibility or discrimination: Understanding jurors' reactions to accurate and inaccurate child and adult eyewitnesses. In G. S. Goodman & B. L. Bottoms (Eds.), *Child victims, child witnesses: Understanding and improving testimony* (pp. 181–197). New York: Guilford Press.
- Lindberg, M. A., Kiefer, J., & Thomas, S. (2000). Eyewitness testimony for physical abuse as a function of personal experience, development, and focus of study. *Journal of Applied Developmental Psychology*, 21(5), 555–591.
- McGough, L. S. (1996). Commentary: Achieving real reform: The case for American interviewing protocols. *Monographs of the Society for Research in Child Development*, 61, Nos. 4–5, 188–203.
- Melton, G. B., Goodman, G. S., Kalichman, S. C., Levine, M., Saywitz, K. J., & Koocher, G. P. (1995). Empirical research on child maltreatment and the law. *Journal of Clinical Child Psychology*, 24(Suppl.), 47–77.
- Memon, A., Bull, R., & Smith, M. (1995). Improving the quality of police interviews: Can training in the use of cognitive techniques help? *Policing Society*, 5, 53–68.
- Memon, A., Wark, L., Holley, A., Bull, R., & Koehnken, G. (1997). Eyewitness performance in cognitive and structured interviews. *Memory*, 5(5), 639–656.
- Orbach, Y., Hershkowitz, I., Lamb, M. E., Sternberg, K. J., Esplin, P. W., & Horowitz, D. Assessing the value of structured protocols for forensic interviews of alleged child abuse victims. *Child Abuse & Neglect*, 24(6), 733–752.
- Ornstein, P. A. (1996). Commentary: To interview a child: Implications of research on children's memory. *Monographs of the Society for Research in Child Development*, 61, Nos. 4–5, 215–222.
- Ornstein, P. A., Baker-Ward, L., Myers, J., Principe, G. F., & Gordon, B. N. (1995). Young children's long-term retention of medical experiences: Implications for testimony. In F. E. Weinert & W. Schneider (Eds.), *Memory performance and competencies: Issues in growth and development* (pp. 349–371). Hillsdale, NJ: Erlbaum.
- Peterson, C. (1999). Children's memory for medical emergencies: 2 years later. *Developmental Psychology*, 35(6), 1493–1506.
- Pezdek, K., & Taylor, J. (2000). Discriminating between accounts of true and false events. In D. F. Bjorklund (Ed.), *False-memory creating in children and adults: Theory, research, and implications* (pp. 69–91). Mahwah, NJ: Erlbaum.

- Pipe, M., Gee, J., Wilson, C., & Egerton, J. M. (1999). Children's recall 1 or 2 years after an event. *Developmental Psychology, 35*(6), 781-789.
- Poole, D. A., & Lamb, M. (1998). *Investigative interviews of children: A guide for helping professionals*. Washington, DC: American Psychological Association.
- Poole, D. A., & Lindsay, D. S. (1998). Assessing the accuracy of young children's reports: Lessons from the investigation of child sexual abuse. *Applied & Preventive Psychology, 7*, 1-26.
- Poole, D. A., & White, L. T. (1991). Effects of question repetition on the eyewitness testimony of children and adults. *Developmental Psychology, 27*(6), 975-986.
- Poole, D. A., & White, L. T. (1995). Tell me again and again: Stability and change in the repeated testimonies of children and adults. In M. S. Zaragoza, J. R. Graham, G. C. N. Hall, R. Hirschman, & Y. S. Ben-Porath (Eds.), *Memory and testimony in the child witness* (pp. 24-43). Thousand Oaks, CA: Sage.
- Raskin, C. C., & Esplin, P. W. (1991). Assessment of children's statements of sexual abuse. In J. Doris (Ed.), *The suggestibility of children's recollections: Implications for eyewitness testimony* (pp. 47-55). Washington, DC: American Psychological Association.
- Rudy, L., & Goodman, G. S. (1991). Effects of participation on children's reports: Implications for children's testimony. *Developmental Psychology, 27*, 527-538.
- Saywitz, K. J. (1995). Improving children's testimony: The question, the answer, and the environment. In M. S. Zaragoza, J. R. Graham, G. C. N. Hall, R. Hirschman, & Y. S. Ben-Porath (Eds.), *Memory and testimony in the child witness* (pp. 113-140). Thousand Oaks, CA: Sage.
- Steller, M. (1989). Recent developments in statement analysis. In J. C. Yuille (Ed.), *Credibility assessment* (pp. 135-154). Dordrecht, The Netherlands: Kluwer.
- Sternberg, K. J., Lamb, M. E., Hershkowitz, I., Esplin, P. W., Redlich, A., & Sunshine, N. (1996). The relation between investigative utterance types and the informativeness of child witnesses. *Journal of Applied Developmental Psychology, 17*, 439-451.
- Sternberg, K. J., Lamb, M. E., Hershkowitz, I., Yudilevitch, L., Orbach, Y., Esplin, P. W., et al. (1997). Effects of introductory style on children's abilities to describe experiences of sexual abuse. *Child Abuse & Neglect, 21*, 1133-1146.
- Sternberg, K. J., Lamb, M. E., Orbach, Y., Esplin, P. W., & Mitchell, S. (2001). Use of a structured investigative protocol enhances young children's responses to free recall prompts in the course of forensic interviews. *Journal of Applied Psychology, 86*(5), 997-1005.
- Stevenson, K. M., Leung, P., & Cheung, K. M. (1992). Competency-based evaluation of interviewing skills in child sexual abuse cases. *Social Work Research & Abstracts, 28*, 11-16.
- Steward, M. S., & Steward, D. S. (1996). Interviewing young children about body touch and handling. *Monographs of the Society for Research in Child Development, 61*, Nos. 4-5.
- Tobey, A. E., Goodman, G. S., Batterman-Faunce, J. M., Orcutt, H. K., & Sachsenmaier, T. (1995). Balancing the rights of children and defendants: Effects of closed-circuit television on children's accuracy and jurors' perceptions. In M. S. Zaragoza, J. R. Graham, G. C. N. Hall, R. Hirschman, & Y. S. Ben-Porath (Eds.), *Memory and testimony in the child witness* (pp. 214-239). Thousand Oaks, CA: Sage.
- Walker, A. G., & Warren, A. R. (1995). The language of the child abuse interview: Asking the questions, understanding the answers. In T. Ney (Ed.), *True and false allegations of child sexual abuse: Assessment and case management* (pp. 153-162). New York: Brunner/Mazel.
- Warren, A. R., & Lane, P. (1995). Effects of timing and type of questioning on eyewitness accuracy and suggestibility. In M. S. Zaragoza, J. R. Graham, G. C. N. Hall, R.

- Hirschman, & Y. S. Ben-Porath (Eds.), *Memory and testimony in the child witness* (pp. 44–60). Thousand Oaks, CA: Sage.
- Warren, A. R., Woodall, C. E., Hunt, H. S., & Perry, N. W. (1996). “It sounds good in theory, but”: Do investigative interviewers follow guidelines based on memory research? *Child Maltreatment, 1*(3), 231–245.
- Warren, A. R., Woodall, C. E., Thomas, M., Nunno, M., Keeney, J., Larson, S., et al. (1999). Assessing the effectiveness of a training program for interviewing child witnesses. *Applied Developmental Science, 3*, 128–135.
- Yuille, J. C., Hunter, R., & Joffe, R. (1990). A coordinated approach to interviewing in child sexual abuse investigations. *Canada's Mental Health, 38*(2/3), 14–17.
- Yuille, J. C., Hunter, R., Joffe, R., & Zaparniuk, J. (1993). Interviewing children in sexual abuse cases. In G. S. Goodman & B. L. Bottoms (Eds.), *Child victims, child witnesses* (pp. 95–115). New York: Guilford Press.

*Received April 20, 2002*