Reliability and Credibility of Young Children's Reports

From Research to Policy and Practice

Maggie Bruck — *McGill University*
Stephen J. Ceci and Helene Hembrooke — *Cornell University*

In this article, some issues, concerns, and research regarding the interviewing of young child witnesses are reviewed. The article focuses on research on suggestibility and the influence of various interviewing techniques on the reliability and credibility of young children's reports. Implications of this research for future research and for policy are discussed.

By the 1980s, there was an enormous change in society's sensitivity to and recognition of the problems of violence and abuse that were suffered by children. Spurred by an increased awareness of the pervasiveness of one crime, child sexual abuse, state after state revised its criminal procedures to enable prosecutors to deal more effectively with victims and defendants. This led to important changes in the legal system not only in the United States but also in other countries in the Western world (see Bottoms & Goodman, 1996; Davies, Lloyd-Bostock, McMurran, & Wilson, 1995). The most important of these changes was the relaxation of standards that had prevented many children from testifying in criminal and civil cases. Children for the first time were allowed to provide uncorroborated testimony in cases concerning sexual abuse—a crime that by its very nature often does not involve an eyewitness other than the perpetrator and the victim. A second major change involved the elimination of the competency requirement for child witnesses. There was also a raft of other legal changes and proposals that had far-reaching consequences for the structure of the legal system and in some cases for the constitutional safeguards of the defendants. In some cases, children are now permitted to give sworn and unsworn testimony behind screens that shield them from the defendant's gaze or by closed circuit TV (*Maryland v. Craig,* 1990). In some states, there has been a broadening of the receipt of children's hearsay statements (e.g., *Idaho v. Wright,* 1990; *White v. Illinois,* 1992; for a description of these cases and other changes, see Goodman et al., 1992; McGough, 1994).

Giving children their day in court has opened a Pandora's box of issues not only for the legal community but also for the mental health, social services, and social sciences communities. These issues have been magnified by the media's attention to a number of cases that either call into question the adequacy of existing safeguards for the protection of abused children (e.g., children, such as Lisa Steinberg, who were beaten to death by an abuser who was previously investigated for suspected abuse) or that call into question existing pretrial and trial procedures used to obtain and evaluate the testimony of young children (e.g., young children, such as those in Kern County, California, who made horrific allegations of sexual abuse after many hours and days of questioning by the police and social service workers; see Nathan & Snedeker, 1995). These two situations reflect two serious problems: underreporting of actual abuse and overreporting of abuse that has not occurred. To effectively deal with these two problems, we require a clear understanding of the strengths and weaknesses of children's memory and children's ability to report experienced events. The present article, while addressing both types of problems, focuses on the weaknesses of young children's reporting and memory. Of specific interest is the accuracy of children's reports and the degree to which they can be positively or negatively influenced by suggestive interviewing techniques. After reviewing the literature in this area, we discuss some of the unique strengths and weaknesses of children's recollections and conclude with the challenge this literature poses for social scientists and policymakers to develop protocols and procedures to maximize children's accuracy with the minimum abridgment of defendants' rights.

Child Sexual Abuse

The Scope of the Problem

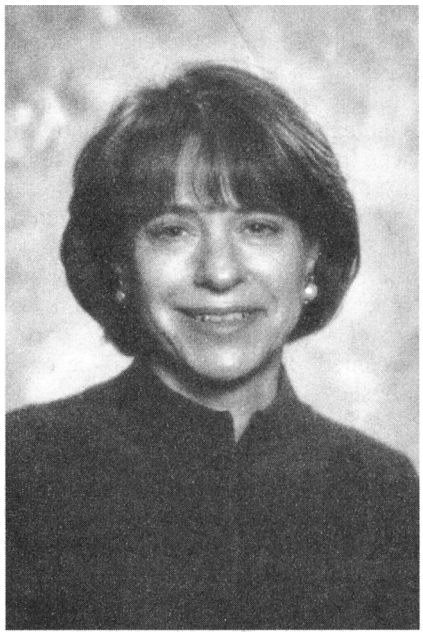
Until very recently, there has been an annual increase in the number of reported and substantiated cases of sexual

Maggie Bruck, Department of Psychology, McGill University, Montreal, Quebec, Canada; Stephen J. Ceci and Helene Hembrooke, Department of Human Development and Family Studies, Cornell University.

We gratefully acknowledge support from the National Science Foundation and the Natural Sciences and Engineering Research Council. We are also grateful to Lucy McGough, Debra Poole, and James Wood for their instructive comments.

Correspondence concerning this article should be addressed to Maggie Bruck, Department of Psychology, McGill University, 1205 Dr. Penfield Avenue, Montreal, Quebec, Canada H3A 1A1. Electronic mail may be sent to bruck@hebb.psych.mcgill.ca.

 0003-066X/98/$2.00
Vol. 53, No. 2, 136–151



Maggie Bruck
Photo by Posen

abuse. The increase in these estimates of abuse in part reflects society's growing sensitivity and hence recognition of existing problems that was prominently marked by the Mondale Act (1974), which mandated the reporting of suspected physical and sexual abuse involving children. The American Humane Association (1988) estimated a 2,000% increase in reports of sexual abuse between 1976 and 1986. In a more recent survey, it was conservatively estimated that between 1986 and 1993 there was an 83% increase in the number of sexually abused children (Sedlak & Broadhurst, 1996). However, for the first time in several decades this trend may now be reversing. In 1994, there were 141,628 substantiated reports of sexual abuse of children under the age of 18; in 1995, this figure was reduced to 126,095 (U.S. Department of Health and Human Services, National Center on Child Abuse and Neglect, 1996, 1997). Similar reductions in the rates of substantiated reports of sexual abuse were recently reported by Wang and Daro (1996). Despite these decreases, an enormous number of children are still sexually abused.

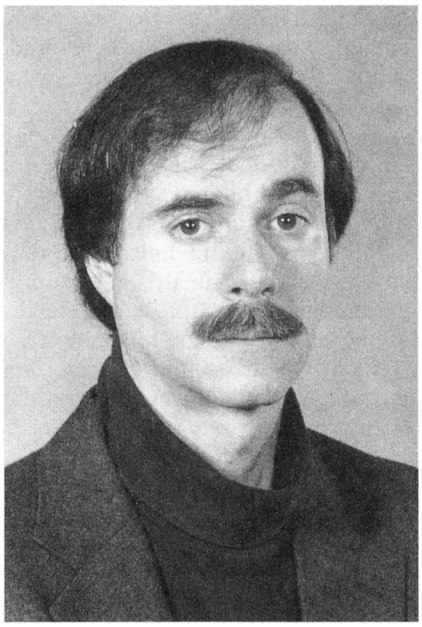
An important issue regarding these figures is the degree to which they represent underreporting or overreporting.[1] On the one hand, it is argued that incidence estimates are very conservative because these include only reported cases; there are many cases of abuse that are never reported. For example, Kalichman (1993) estimated that only 40% of maltreatment cases are reported. The concerns about overreporting focus on the number of substantiated cases that are incorrectly classified. Some of these cases may be the result of false allegations, whereas others may be simply the result of false suspicions. Estimates of false-positive cases range from 5% to 35% (see Ceci & Bruck, 1995; Poole & Lindsay, in press). At times, some of these misclassified cases are detected and resolved. However, at other times, these may end up in the courtroom where children make false accusations about abuse perpetrated by their teachers, babysitters, or parents.

Concerns about underreporting and overreporting of child sexual abuse have been the impetus for a wave of research on children's autobiographical memory, children's suggestibility, and adults' judgments regarding the credibility of child witnesses. Some have argued that the surge of interest in the suggestibility and reliability of children's reports represents a step backward to the views of the earlier part of this century, when the prevailing wisdom was that children were dangerously vulnerable to coaching and suggestions and thus should not be admitted as courtroom witnesses (see Ceci & Bruck, 1993). However, the intent or impact of the new wave of research described below was not to discredit sexual abuse allegations made by children, but merely to set up a context in which the allegations could be realistically evaluated. From a scientific point of view, this recent research has led to a greater understanding of the strengths and weaknesses of children's cognitive and social development. It has led us away from the strong positions that (a) children are hypersuggestible sponges, incapable of accurately remembering and reporting events, or (b) children have the same cognitive structures and mechanisms as adults and are as resistant to suggestions and as able to remember and report events as adults. On the applied side, this newer research is beginning to have some impact on the legal system in terms of the decisions that are made by trial and appellate courts (e.g., *State v. Michaels,* 1994; *United States v. Rouse,* 1996).

Diagnosing Child Sexual Abuse

Although children testify in a range of criminal and civil cases, the research on children's suggestibility has had its largest impact in cases involving allegations of sexual abuse. This is primarily because of the lack of scientifically validated criteria for the diagnosis of sexual abuse. Specifically, in most cases there are no medical findings considered to be sufficiently diagnostic to substantiate abuse (see Bays & Chadwick, 1993). Usually there is no medical evidence, either because of lengthy delays in reporting or, as is often the case, because there is no penetration (e.g., fondling, exhibitionism, or oral copulation are the most common crimes). In the statistically rare case where genital or anal abnormalities are found, similar abnormalities can sometimes be found among nonabused children (Berenson, Heger, & Andrews, 1991). And if the problems associated with lack of hard medical evidence were not bad enough, there does not appear to be a single psychological profile that is diagnostic of child sexual abuse (Kendall-Tackett, Williams, & Finkelhor, 1993). Although there are a number of symptoms associated with validated cases of abuse, it turns out that

[1] Elsewhere, it has been surmised that the number of unreported cases is far greater than the number of falsely reported ones (Ceci & Bruck, 1995), although this is speculation.



Stephen J. Ceci
Photo by Charles Harrington, Cornell University Photography

either these symptoms are common childhood problems (e.g., regressive toiletting, acting out, night tremors) or these symptoms are behavioral problems that are commonly found in other childhood psychopathologies. Consequently, to diagnose child sexual abuse, one must rely on the verbal report of the child witness.

In evaluating the child's testimony, it has become clear that it is of primary importance to understand the evolution of the child's reports. The following pattern, one that is found in many cases involving child sexual abuse, has raised the most concerns. Here the child is initially silent: She does not make any unsolicited or spontaneous statements about abusive acts. Rather, the allegations emerge once an adult has a suspicion that something has occurred and starts to question the child. At first, the child denies the event happened, but with repeated questioning, interviewing, or therapy, the child may eventually make a disclosure. Sometimes after the disclosure is made, the child may recant, only to later restate the original allegation. There are two different interpretations of this pattern.

The first interpretation is that the progression from silence to denial to disclosure to recanting to restatement is common and perhaps even diagnostic of sexual abuse. Some professionals claim that children have a great deal of difficulty disclosing and may later recant because they are afraid or ashamed or even believe themselves to be culpable (for a review, see Bradley & Wood, 1996). Although there are some formal models of the disclosure process (e.g., Summit's 1983 child sexual abuse accommodation syndrome includes a series of stages: secrecy, helplessness, entrapment, disclosure, and retraction), it is important to point out that these models were not derived from scientific studies but from clinical intuitions. Nonetheless, there are a few studies that support the view that when directly asked about abuse, it is common for sexually abused children to not readily or consistently disclose their abuse. The most frequently cited of these studies was conducted by Sorensen and Snow (1991), who examined 116 cases from a sample of 630 children who had received therapy for sexual abuse. For the majority of the children, the disclosures were accidental, and at some point 75% of the children had denied that abuse had occurred. Even after making a disclosure, 22% of the children recanted their previous disclosures. In another study, Gonzalez, Waterman, Kelly, McCord and Oliveri (1993) found that 27% of 63 children in therapy recanted disclosure of day-care sexual and ritualistic abuse.

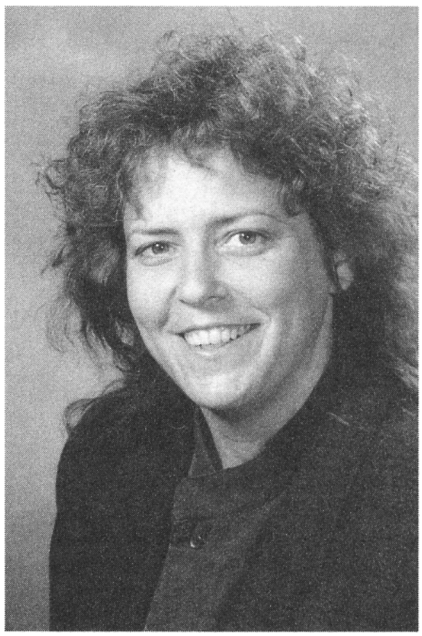
Unfortunately, there are a number of methodological features that mar the interpretation of such data. Of primary concern is the possibility that a number of the children in these studies may not have been sexually abused and that their disclosures were the result of the interviewing and therapeutic process (see Ceci & Bruck, 1995). This is particularly salient in the Gonzalez et al. (1993) study, which examined the disclosure patterns of the McMartin Preschool child witnesses; claims of ritualistic abuse by the McMartin children were never substantiated, and they appear not to have been accepted as credible by jurors.[2]

There are other studies that provide a different perspective on the process of disclosure. For example, Jones and McGraw (1987) found only an 8% recantation rate among 309 validated sexual abuse cases seen at a child protection agency. In the most recent study of the patterns of disclosure among sexually abused children, Bradley and Wood (1996) found that among 234 validated cases, 5% of the children denied the abuse, and only 3% recanted their earlier reports of abuse. Clearly, although a small percentage of youngsters do appear to disclose their abuse reluctantly, with a smaller percentage subsequently recanting their disclosures, the overwhelming majority of children appear to maintain their claims and never deny them to officials once they are questioned.[3]

Despite the frail empirical evidence for the prevalence of the disclosure process among sexually abused children, it may well live on as a stubborn urban legend among frontline workers. For example, we recently surveyed the opinions and beliefs of 26 highly trained child protection workers. These investigators not only stated that recantation was part of the normal process of disclo-

[2] In general, claims of ritualistic abuse (e.g., animal or human sacrifices, cannabilism, witchcraft) that have been made by a number of child witnesses (for details, see Ceci & Bruck, 1995; Nathan & Snedeker, 1995) are viewed skeptically by the scientific community. To date, there is an overwhelming lack of evidence from two national surveys (Bottoms, Shaver, & Goodman, 1996; La Fontaine, 1994) for these claims.

[3] It is beyond the scope of this article to explain discrepancies among studies on disclosure patterns or to postulate the factors that make some children reticent to disclose, such as the investigative setting, the supportiveness of the caretaker, and the absence of disclosure previous to the investigative interviews (see Sternberg et al., 1996).



Helene Hembrooke
Photo by Charles Harrington, Cornell University Photography

sure, but they also indicated that the research strongly supported their belief. This could translate into the following belief among child protective service workers: Abused children must be pursued or they will never disclose their abuse, and one should not readily accept their denials or recantations because truly abused children usually show these very behaviors (Conerly, 1986; MacFarlane & Krebs, 1986).

There is a second interpretation of a disclosure pattern that begins with silence, then progresses to denial, and eventuates in disclosure (and sometimes ends in recantation): For some children, the disclosures may reflect the use of suggestive interviewing techniques, which, in some situations, may elicit false reports. To discuss this hypothesis, we first focus on the concept of *suggestive interviewing techniques.*

Suggestive Interviews

Before the 1980s, most studies of suggestibility involved asking children a misleading question (i.e., a question that contains a false supposition) about some experienced or observed event (e.g., a story, a school demonstration). A consistent finding of this literature was that younger children were more suggestible than older children (for a review, see Ceci & Bruck, 1993). However, for the following reasons, this literature was of little value in assessing issues of reliability or suggestibility of children who make allegations of sexual abuse or other potentially distressing events. First, the age of the children studied was problematic. There was but one study in the first 80 years of this century that included preschool children (Lipmann & Wendriner, 1906)—the very age group that is the cause of most concern for modern-day courts. That is, a disproportionate number of sexually abused children are preschoolers, and a disproportionate number of court cases involve preschool witnesses (see Ceci & Bruck, 1995). Second, the children in these studies were questioned about neutral events that had little personal salience. For example, Lipmann and Wendriner (1906) found that preschoolers were progressively susceptible as the strength of misleading questions was increased. Thus, when four- to six-year-olds were asked about a nonexistent cabinet, only 6% falsely assented to the question, "Is there a cabinet in the room?"; when asked, "Isn't there a cabinet in the room?" the false assent rate rose to 25%. Finally, false answers reached a maximum of 56% when children were asked, "Is the door open in the cabinet in the room?" A related point that is illustrated by the previous example is that the questioning of the children in the experimental settings seemed to bear little if any similarity to the conditions under which children are questioned in actual cases.

In actual investigations, because children are rarely questioned under such neutral conditions about relatively benign events, it became clear to social scientists that there would have to be major revisions to existing paradigms and concepts to provide pertinent information to the court about whether or not a child's testimony could be the product of the interviewing methods. There have been three important changes in the direction of this research. First, preschool children were included in many of these newer studies. Second, studies increasingly were designed to examine children's suggestibility about events that were personally salient, that involved bodily touching, or that involved insinuations of sexual abuse. Third, the concept of suggestive techniques was expanded from the traditional view of asking misleading questions or planting misinformation to using a larger range of interviewing devices that will be discussed below. A recurring theme of these newer studies is the attempt to question children about the main actions that occurred during the experienced event rather than only about the peripheral details, such as the color of an actor's shoes. The ultimate challenge has been to ask questions in an ethically permissible manner about whether or not sexual actions occurred during these events.

With great ingenuity, a number of researchers have met this challenge. For example, Saywitz, Goodman, Nicholas, and Moan (1991) questioned five- to seven-year-old girls about the details of a medical examination that for some children included a genital examination. The children were asked open-ended, direct, and misleading questions about touching and were also asked to demonstrate what happened to them by using anatomically detailed dolls. In a series of studies, Ornstein and his colleagues (Ornstein, Baker-Ward, Myers, Principe, & Gordon, 1995) asked three- to seven-year-old children about their annual pediatric visits. These studies assessed the rate at which memories fade over different periods of delay as well as the degree to which children falsely include nonoccurring events as part of their reports.

Thus, results of the Saywitz et al. (1991) and Ornstein et al. (1995) studies provide data as to how accurately children report salient events, which may in-

clude bodily contact, when they are questioned immediately following the events or after up to a three-month delay. These data also reflect how accurately children respond to open-ended and both direct questions and misleading questions when they are questioned by a neutral, unbiased interviewer. Under these conditions, the children were fairly but not entirely accurate about a number of salient events that involved bodily touching. The results also show a typical pattern, found across many studies, in which children provide more information in response to specific compared with open-ended questions. For example, in the Saywitz et al. study, few children mentioned the genital examination in response to the question, "Tell me everything that happened," but many did provide the required information in response to more specific questions. However, although children generally provide more information to specific questions, it is generally the case that overall, accuracy rates are higher for responses to open-ended questions. Furthermore, accuracy of responses to specific and misleading questions increases as a function of age (with preschoolers being the least accurate) and as a function of the delay between the interview and the actual event. Some data indicate that when accuracy drops off, it is not merely the case that children forget and therefore make errors of *omission* (i.e., failing to recall an actual event), but they also make errors of *commission* (falsely claiming to have experienced a nonevent). In Ornstein et al.'s study, children, especially the younger children, reported events that never happened. These reported nonevents included not only acts that could conceivably occur in a doctor's office but also acts that would not occur in the doctor's office and that would have connotations of abuse, at least to some adults (e.g., "Did the doctor lick your knee?" "Did the nurse sit on top of you?").

Although these data are important for understanding how specific types of questions may alter children's patterns of reporting, recall, and forgetting, they may have limited value for assessing the accuracy of children's statements when these are obtained under more suggestive situations that occur in some criminal investigations. In contrast to a number of child witnesses who are caught up in actual investigations, the children in these research studies were not repeatedly interviewed about alleged abusive events; identical questions were not repeated within and across interviews; nor were there threats or inducements to have the participants reply in a certain way. The interviewers in these studies were supportive and neutral, in contrast to what can happen when anxious parents, therapists, and legal officials, often repeatedly and over lengthy intervals, question child witnesses. Because these important elements are missing from research studies such as those cited above, it makes it difficult to generalize to forensic situations. It is this concern that has motivated another group of modern researchers to look more closely at the structure of conversations and interviews[4] between children and adults and to examine the effects of various interviewing practices on the accuracy of children's reports. As we show below, the development of this approach has been particularly important in understanding the putative pattern of disclosure that begins with secrecy and denial and that eventuates in allegations. In the next section, we discuss some suggestive elements of interviews that take place between adults and children.

Interviewer Bias and Suggestive Interviewing Techniques

A major dimension along which interviewers can be characterized is that of *interviewer bias*. Interviewer bias characterizes those interviewers who hold a priori beliefs about the occurrence or nonoccurrence of certain events and, as a result, mold the interview to elicit statements from the interviewee that are consistent with these prior beliefs. One of the hallmarks of interviewer bias is the single-minded attempt to gather only confirmatory evidence and to avoid all avenues that may produce negative or inconsistent evidence. Thus, while gathering evidence to support their hypotheses, interviewers may fail to gather any evidence that could potentially disconfirm their hypotheses. Biased interviewers do not ask questions that might provide alternate explanations for the allegations (e.g., "Did your mommy and daddy tell you that this happened, or did you see it happen?"). Nor do biased interviewers ask the child about events that are inconsistent with their hypotheses (e.g., "Who else beside your teacher touched your private parts? Did your mommy touch them, too?"). And biased interviewers do not challenge the authenticity of the child's report when it is consistent with their hypotheses (e.g., "It's important to tell me only what you saw, not what someone may have told you," or "Did that really happen?"). When children provide inconsistent or bizarre evidence, it is either ignored or else interpreted within the framework of the biased interviewer's initial hypothesis. In short, interviewer bias can be found whenever interviewers think they know the answers before the child divulges them and whenever interviewers view their task as one of "getting the goods" on the defendant.

Interviewer bias influences the entire architecture of interviews, and it is revealed through a number of different component features that are suggestive. For example, to obtain confirmation of their suspicions, biased interviewers may not ask children open-ended questions such as "What happened?" but may quickly resort to a barrage of very specific questions, many of which are repeated, and many of which are leading in the sense that the question stem presupposes the desired answer. When interviewers do not obtain information that is consistent

[4] Although interviews may be highly structured, they need not be. An interview, at minimum, is a verbal interaction between at least two people in which one of the participants (the interviewer) has the goal of obtaining specific information from one of the participants (the interviewee). As such, interviews are a particular type of conversation that can be carried out by a wide variety of professionals and nonprofessionals, such as child protection workers, police officers, mental health professionals, attorneys, parents, or teachers.

with their suspicions, they may repeatedly interview children until they do obtain such information, sometimes subtly reinforcing responses consistent with their beliefs. Thus, child witnesses are often interviewed over a prolonged period of time, and they are reinterviewed on many occasions about the same set of suspected events (for a review, see Ceci, Bruck, & Rosenthal, 1995).

Stereotype inducement is another strategy that is sometimes found in biased interviews with children. Interviewers using this strategy give the children information about some characteristic of the suspected perpetrator. For example, children may be told that a person who is suspected of some crime "is bad" or "does bad things." For example, in one case, a preschooler told her interviewer that she was glad that the defendant was in jail because he was bad. When asked why she thought that he was bad, the child replied, "My mom told me."

Interviewer bias is also reflected in the atmosphere of the interview. Sometimes, interviewers provide much encouragement during the interview to put children at ease and to provide a highly supportive environment. Such encouraging statements can, however, quickly lose their impartial tone if an interviewer *selectively* reinforces children's responses by positively acknowledging statements (e.g., through the use of vigorous head nodding, smiling, and statements such as "Wow, that's great!") that are consistent with the interviewer's beliefs or hypotheses or by ignoring other statements that do not support the interviewer's beliefs. Some interviewers who feel an urgency and responsibility to obtain the desired disclosure may even use threats and bribes. For example, in some cases, children have been told that they will be reunited with their parents if they just tell the investigators how they were abused by their parents. To obtain full compliance from the children, interviewers often try to engage the children by co-opting their cooperation by telling them that they are helpers in an important legal investigation (in one case, children were given plastic police badges and taken on a visit to the police station) and sometimes by telling the children that their friends have helped or already told and that they should also tell. The following blatant example is taken from the Kelly Michaels case (for other examples, see Ceci & Bruck, 1995):

> Interviewer: Do you know that I've been talking to a lot of your buddies? I've been talking to Kit and I've been talking to Mart and . . . we've been talking about some stuff that's not so nice that's been happening at school with Kelly. . . . You can help us with giving some information on your friends that were hurt, okay?

Interview bias is also reflected in the use of some techniques that are specific to interviews between professionals and children. One of these involves the use of anatomically detailed dolls in investigations of sexual abuse. It is thought that these props facilitate reports of sexual abuse for children with limited language skills, for children who feel shame and embarrassment, and for children with poor memories of the abusive incident. In some cases, when interviewers suspected abuse before the children had made any allegations, they gave children anatomically detailed dolls and asked them to show how they had been sexually abused.

Another professional technique involves *guided imagery* or *memory work.* Interviewers sometimes ask children to first try to remember or pretend if a certain event occurred and then to create a mental picture of the event and to think about its details. In some cases, interviewers asked children to pretend about events that were consistent with the interviewers' beliefs but that had not been reported by the child witnesses.

This description of the architecture of interviews is based on our analysis of transcripts that have been made available to us by judges, attorneys, parents, and medical and mental health professionals who have concerns about the conduct of an interview and its potential impact on the reliability of a child's reports. As such, our description may not be representative of many or even most of the interviews carried out with children in forensic or therapeutic situations: Undoubtedly, there are many interviews that do not contain any of the features in our description. However, there is some recent evidence to indicate that a number of interviewers do in fact use some of the techniques described (Bull & Cherryman, 1995; Hulse, 1994; Lamb et al., 1996; Warren, Woodall, Hunt, & Perry, 1996; Yuille, Marxsen, & Menard, 1993). In these studies, interviewers mainly relied on specific or leading questions; several times during the interviews, they introduced information that the children had not volunteered, and they frequently repeated that new information in the course of a single interview.

Having qualified the basis of our descriptions, it is also important to state that the research in this area extends beyond a mere cataloging of the elements of potential bias. The next step, the empirical step, is to assess how these various biased components, in isolation as well as in combination, influence the accuracy of children's reports. This framework for describing and examining the components of interview bias greatly expands the meaning of the term *suggestive.* In the first 80 years of this century, most of the research on suggestibility focused on the effects of asking a single misleading question or of providing erroneous postevent information on the subsequent accuracy of reports. Now, the study of suggestive techniques includes strategies of repeated interviewing, repeated questioning within interviews, the use of threats and rewards, and other techniques that are intended to put the child at ease and to facilitate disclosure. Suggestive interviews now are conceived of as a complex comingling of motives, threats, and inducements, which may appear in the form of misleading questions, but not always.

A number of recent studies have examined the influence of a number of the biased techniques just described (for a review, see Ceci & Bruck, 1995). In these studies, children typically participated in some event and then were interviewed by using one or more of the alleged

suggestive interviewing techniques listed above. The results of these newer studies show that the suggestive interviewing techniques just described can compromise the accuracy of children's reports. When used in combination, such techniques can be especially detrimental to the accuracy of young children's reports. In some of these studies, children fabricated whole events that never occurred; at times they were led to confuse suggestions with actual events to the point that they later insisted that the suggested events actually occurred. Importantly, these suggestive techniques do not merely influence recall of peripheral, unimportant details, but they lead to false claims about a wide range of events, many of which are personally meaningful, such as bodily touching that could be interpreted by adult interviewers as sexual in nature. And because children make false claims about unpleasant events, these newer studies are not open to the older criticism that children are only suggestible about unimportant, neutral details (Melton, 1992; Melton & Thompson, 1987).

For example, studies examined three-year-old (Bruck, Ceci, Francoeur, & Renick, 1995) and four-year-old (Bruck, Ceci, & Francoeur, 1995) children's memories of a just-completed medical examination in which half of the children received a genital examination and half did not. After the examination, the children were given an anatomically detailed doll and were asked to show on the doll how the doctor had touched them. In addition, they were provided with props (e.g., a spoon, a stethoscope) and asked to show how the doctor had used these. The wording of some of the questions was misleading for the children who had not received a genital examination, although these same questions could be considered to be correctly leading for children who *had* received a genital examination (e.g., "Show me on the doll how the doctor touched your penis").

The results for the three- and four-year-old children were almost identical. A significant proportion of children showed inaccurate touching. Specifically, a substantial number of children who had a genital examination failed to show accurate touching when touching had in fact occurred. Some of these errors were omission errors (the children did not show any touching), but for the girls most of the errors involved showing overtouching (they inserted fingers into the genitalia or anus when this had not occurred). At the same time, a significant proportion of children who had not had a genital examination showed genital touching on the doll when it had *not* happened, a form of commission error. In the latter situation, had the child been touched in the manner that was demonstrated, it might have been quite painful. That is, a number of children (especially the girls) inaccurately showed the pediatrician inserting his finger or props into their genitalia or buttocks.

It seems that there are similar concerns about the use of anatomically detailed dolls with children as old as six years of age. Steward and Steward (1996) interviewed children (ages three to six years) three times after a pediatric clinic visit. With each interview, children's false reports of anal touching increased; by the final interview, which took place six months after the initial visit, more than one third of the children made such errors of commission (see also Rawls, 1996, for errors of commission by children interviewed using a diagram of body parts).

Other studies reveal how the combination of certain suggestive interviewing techniques can produce false reports not only about touching but also about emotions. For example, in one study six-year-olds were suggestively and repeatedly interviewed about some salient characteristics of a medical examination that had occurred approximately 12 months previously (Bruck, Ceci, Francoeur, & Barr, 1995). During this medical examination, the children had received an inoculation that for many was stressful and that for most was painful during the next few days. As a result of our repeated use of misinformation a year after the procedure, a number of children eventually came to report that our female research assistant rather than their male pediatrician had actually given them their inoculation and an oral vaccine. The results of this study highlight some other deleterious effects of repeating misinformation across interviews: Not only did some children directly incorporate the misinformation into their subsequent reports, but they also made other inaccurate claims that were not suggested (e.g., falsely reporting that the female research assistant had checked their ears and nose). These statements are inferences that are consistent with the erroneous suggestion that the female research assistant had administered the shot: therefore she must have been the doctor, and therefore she carried out other procedures commonly performed by doctors. Finally, compared with control children who were not given any suggestions about how much the shot hurt or how much they cried, the misinformed children routinely underestimated their level of pain and crying as a result of erroneous suggestions about how brave and courageous they had been. Thus, young children may use suggestions in highly productive ways to reconstruct and at times distort reality about unpleasant bodily events.

These medical setting studies do not reveal whether it is easier to influence children's reports for unpleasant than for pleasant events. Data from two other studies do address this concern, however. In the first study (Ceci, Loftus, Leichtman, & Bruck, 1994), parents of preschool children told researchers about four true events involving their preschool children: a pleasant event (e.g., a birthday party), an unpleasant event (e.g., death of a pet), and two neutral events (e.g., wearing a blue sweater to school). The parents also verified that a list of four false events, devised by the research team, had never occurred to their children. There was a false pleasant event (taking a ride in a hot-air balloon), a false unpleasant event (falling off a bike and getting stitches) and two neutral events (waiting for a bus and watching a friend waiting for a bus). Preschool participants were asked to create images (using visualization techniques) about these real and fictional events and to tell researchers if they actually experienced the imaged events. The children were asked to do this once per week for 11 consecutive weeks. Assents to im-

ages of true events were at ceiling, indicating highly accurate memory of actual experiences. With time, however, children increasingly assented to fictional events, but their rates of assent for the pleasant event (a ride in a hot-air balloon) were much higher than those for the unpleasant event (a bicycle accident). These results indicate that it may be easier to influence children's reports for pleasant than for unpleasant events but that the latter is not impossible—in fact, it occurred with some frequency.

In a recently completed study (Bruck, Ceci, & Hembrooke, 1997), we compared children's assents to true and false events that were scripted to have positive or negative outcomes. Preschool children were asked to tell us about two true events and about two false events. Each child participated in one of the true events that was staged at their school. This involved the child helping a visitor in the school who had tripped and hurt her ankle. This event was labeled as *true-positive* because the child helped the visitor and was rewarded with verbal praise. The second true event varied across all the children: It involved an actual recent incident for which the child had been punished by the teacher or the parent. This event was labeled as *true-negative.* The *false-positive* event involved helping a lady find her monkey, which had become lost in the park. The *false-negative* event involved witnessing a man steal food from the day care. Unlike the other events, the false-negative event is a criminal act, and thus there are potentially serious implications of assenting to such scenarios.

Children were interviewed on five different occasions about the four events. In the first interview, the children were asked if the event had happened and, if so, to provide as many details as possible about its occurrence. The next three interviews included a combination of suggestive interviewing techniques that have been shown to increase children's assents to false events. These techniques included the use of peer pressure ("Megan and Shonda were there and they told me you were there, too"), visualization techniques ("Try to think about what might have happened"), repeating (mis)information, and providing selective reinforcement. The same interviewer questioned the children for the first four interviews. In the fifth interview, a new interviewer questioned each child about each event in a nonsuggestive manner.

Across the five interviews, all children consistently assented to the true-positive event. However, children were at first reluctant to talk about the true-negative (punishment) event; many of the children denied that the punishment had occurred. With repeated suggestive interviews, the children agreed that the punishment had occurred. Similar patterns of disclosure occurred for the false events; that is, children initially denied the false events, but with repeated suggestive interviews they began to assent to these events. By the third interview, most children had assented to all true and false events. This pattern continued to the end of the experiment.

To summarize, a variety of interviewing techniques can result in young children making false allegations about a wide range of events. Sometimes these false reports involve their own bodies, and sometimes these reports involve false accusations about nonexistent crimes. Some commentators frequently object to these conclusions by claiming that the interviewing conditions used in the suggestibility studies are rarely used by professionals. We have already presented some evidence to dispute this claim. But even if we are wrong, and only a small number of professionals use these techniques, the lesson from scientific studies still stands: It is dangerous to use a number of these techniques. But there is another dimension of interviewing that we have not covered, namely, the use of suggestive techniques by nonprofessionals, such as parents. In a series of studies, Poole and Lindsay (1996, in press) have shown that mild suggestive techniques that are repeated by parents in the context of reading a book to their children can result in substantial memory distortion among children three to eight years of age. Thus, our discussion of the negative impact of suggestive interviewing techniques is not limited to interactions between children and professionals.

The results of our newest study (Bruck et al., 1997) from the monkey and thief scenarios present a more complex picture of children's disclosures and also present a dilemma to professional interviewers. Specifically, the results illustrate the beneficial as well as the harmful consequences of using suggestive techniques to elicit reports from young children. For children who may not want to talk about unpleasant but true events, the use of repeated interviews with suggestive components did prompt them to correctly assent to previously denied events. However, the use of these very same techniques promoted children to assent to events that never occurred. These results provide support for the hypothesis that patterns of disclosure that begin with secrecy and denial and that eventuate in disclosure may reflect the use of suggesting interviewing techniques regardless of whether the reported event did or did not occur. At the end of this article, we discuss how policymakers may find effective ways to balance the risk that interviewers will fail to identify cases of genuine abuse against the risk of pursuing and creating false allegations.

The Credibility of Young Children's Reports

The ability of children to provide accurate reports is often discussed in terms of the credibility of their account or the reliability of their memories. These two terms are not the same, and one does not necessarily imply the other. The term *reliability* simply refers to the accuracy of the report. *Credibility,* on the other hand, refers to the believability that one assigns to a witness's testimony. A judgment about credibility is a subjective reality—it is an individual decision that is not necessarily based on the reliability or consistency of a child's recall but rather on its apparent plausibility. In fact, as we argue below, credibility and reliability can be orthogonal dimensions: Reports that are highly reliable (and accurate) may be judged as not very credible, and conversely, inaccurate reports are sometimes judged as highly credible.

Thus, returning to the research we have reviewed, it could be argued that although the accuracy of children's reports may be negatively influenced by a number of suggestive influences, this does not necessarily mean that they will appear credible to others. An intriguing question is whether a juror, a child development researcher, a child therapist, an experienced social worker, or a judge can differentiate children whose reports are accurate from those whose reports are a product of suggestive interviews. The existing evidence suggests that trained professionals cannot reliably tell the difference between these two kinds of children when the children have been subjected to repeated suggestive interviewing techniques that have been conducted over long periods.

One type of evidence for the above claim is provided by the Sam Stone study (see Leichtman & Ceci, 1995, for details). Here, young children between the ages of three and six were interviewed under a number of different suggestive conditions about a stranger named Sam Stone. The experimenter told some children that Sam Stone was a friend and that he was very clumsy. Over the next few weeks, these children were told numerous stories of Sam Stone's clumsiness. (This technique is called *stereotype induction.*) All children in the experiment eventually did meet Sam Stone: He made one visit to their classroom and was introduced to them during story time. The next day, the teacher showed all the children a soiled teddy bear and a ripped book (Sam had not defiled any of these objects). Every two weeks for the next few months some of the children were provided with repeated misinformation about Sam's visit (they were asked misleading questions, e.g., "When Sam Stone tore the book, did he do it on purpose or was he being silly?"). Finally, at the end of 12 weeks all children were questioned by a new interviewer about what actually happened during Sam Stone's visit.

There were several important findings. First, children who had been repeatedly interviewed with the combination of stereotype induction and repeated misleading questions made the most false reports about what happened when Sam Stone visited the classroom. Second, there were significant age differences: Compared with five- and six-year-old children, three- and four-year-old children were more likely to make false claims about Sam vandalizing the bear and the book, to claim that they saw Sam do these things, and to maintain these claims when challenged by the interviewer.

Third, some children went beyond simple assents to Sam's misdeeds. These children provided false perceptual details as well as nonverbal gestures to embellish their false stories. For example, children used their hands to show how Sam had purportedly thrown a book up in the air; children reported seeing Sam in the playground, on his way to the store to buy chocolate ice cream, or in the bathroom soaking the teddy bear in water before smearing it with a crayon.

Finally, and most important to the issue of credibility, when experts (who included mental health professionals, research psychologists, judges, social workers, and prosecutors) were shown videotapes of children in the Sam Stone study and asked to judge the children's credibility, they were very inaccurate. They judged children whose reports were a product of suggestive interviewing as highly credible and believable, whereas those whose reports were more accurate were often judged to be less believable and credible (see Ceci, Crotteau-Huffman, Smith, & Loftus, 1994; Ceci, Loftus, et al., 1994, for similar results, and see Finlayson & Koocher, 1991; Horner, Guyer, & Kalter, 1993a, 1993b, for variability on professionals' judgments about children suspected of being abused).

Perhaps professionals have difficulty distinguishing accurate from inaccurate reports under conditions of repeated suggestions because many of the children have come to believe what they are telling adults. (We have referred to this state as *false belief* to distinguish it from acts of lying.) Children are not aware that their reports are factually false, and therefore there are no signs of duping, tricking, or rusing. They appear to be motivated to tell the truth; their reports seem quite consistent, embellished, and cohesive. When children believe what they are saying, it can be very difficult to detect errors.

In our recent work (Bruck et al., 1997), we have gone further in our search for potential markers of accurate and inaccurate reports. We hypothesized that perhaps on closer inspection, there may be linguistic markers that differentiate between true and false narratives that are the result of suggestive interviews. For example, we hypothesized that more subtle aspects of children's narratives, such as the total amount of information provided by the children, the number of spontaneous unprompted statements, the cohesiveness of the narrative, or the degree of elaboration, may differentiate true from false narratives. In addition, we examined the evolution of children's narratives to determine if true and false narratives became more similar with repeated interviewing. We also examined the consistency as well as the inconsistency of children's reports across repeated retellings. We reasoned that children may be more likely to repeat the same details and less likely to contradict themselves when repeating true rather than false narratives.

Our selection of linguistic markers was motivated by several areas of research. First, a number of studies have indicated that when children's reports are spontaneous or in response to open-ended questions (e.g., "Tell me what happened"), their reports are more accurate than those prompted by specific questions. (This finding has been consistently reported since the beginning of this century. For an early study, see Stern, 1910, as reported in Ceci and Bruck, 1995. For a recent study, see Peterson and Bell, 1996). Thus, we predicted that the narratives of the true events would contain more spontaneous statements than the narratives of the false events. Second, we selected several measures of narrative coherence (e.g., dialogue, complex temporal markers, elaborations) because these are characteristics of good narratives and more specifically of children's autobiographical narratives (Fivush, Haden, & Adam, 1995). We anticipated

that measures of narrative coherence may occur less frequently in children's false narratives. Third, we examined consistency and inconsistency across interviews, because consistency of a child's report is often one of the most important criteria used by professionals in evaluating the reliability of children's allegations of abuse (Conte, Sorenson, Fogarty, & Rosa, 1991), and inconsistency in young children's reports lowers their credibility in the eyes of mock jurors (Leippe, Romanczyk, & Manion, 1991).

To examine the distinguishability of true and false reports, we used the narratives that children produced in the monkey–thief study described above. The narratives that we examined were those in which the children fully assented to having participated in or witnessed the false or true event. Except for the true-positive event, the children provided few if any details at the first interview. However, by the second interview, there were no significant differences in the amount of information (i.e., the total number of details) that the children provided for the two false events and for the true-positive event. This pattern continued into the final interview. Most of the details were provided spontaneously, and by the second interview, the number of spontaneous utterances were similar for the false and true-positive events. In these analyses and others reported below, the children provided the fewest utterances for the true-negative (punishment) event. There are two possible explanations. First, the children may have provided minimal information because they were uncomfortable and ashamed. Second, the punishment events were less complex (e.g., the child was put in time-out for talking) than the other events, and thus there were fewer actual details to relate.

One of the motivations for repeating interviews, especially with young children, is to provide them with an opportunity to remember important details that they had not originally reported in the first interview. We found that repeated interviewing resulted in reports of new details (called reminiscences), but more reminiscences were produced for false than for true events. It is in fact possible that high rates of reminiscing may signify unreliable reporting, and this may be especially so when the reminiscences are produced after some lengthy delay. For example, Salmon and Pipe (1997) exposed children to a quasi-medical examination and then interviewed them soon after and one year later. They found that children's reminiscences for an event that had been experienced a year previously were largely inaccurate. More research on this important topic is needed before we can confidently state what types and rates of reminiscences are symptomatic of false accounts.

Our examination of the markers of narrative coherence (e.g., simple temporal markers, complex temporal markers, dialogue) and of elaboration (adjectives, adverbs) revealed similar results. After the first interview, there were rarely any between-event differences on these measures—for example, children used dialogue as frequently in true stories as in false stories. When there were differences, the false stories were more coherent than the true stories. Specifically, with time, narratives became more elaborate (the children used more emotional terms, and they included more adjectives and adverbs), but this held mainly for the false stories so that at the third interview, false narratives contained more elaborate details than did true narratives.

Next, we examined the consistency as well as the inconsistency of the children's stories. Consistency refers to the child's mentioning the same detail in more than one interview. Consistency did differentiate true from false stories, with true stories containing more consistent details. The inconsistency analysis did not produce the same results. Inconsistency was defined as the children reporting "A" in one interview but "NOT-A" in another interview. For example, in one interview the child might say that he was with his parents, whereas in a later interview, he might say that he was by himself. The rates of inconsistency were generally low. Importantly, the proportions of inconsistent details were similar across true and false events.

Finally, we examined the number of inaccurate details provided in the true-positive narratives. We did not include minor incorrect details (such as the errors about the appearance of the confederate) but only inaccurate details that would be of major significance in the interpretation of the event. For example, after the first interview some children included the following types of inaccurate details in their narratives: "I called 9-1-1," "She fell many times," and "A man came and helped her." We found that with repeated interviewing, children included more inaccurate details in their narratives. Although the overall number of inaccurate details was relatively small compared with the number of accurate details, these would have a major consequence had they really happened.

The results of this study indicate that it is the first narrative that was elicited by nonsuggestive techniques that allowed the clearest differentiation between true and false stories. This is because children mainly denied the false stories, which as a result contained few details during the first narrative. However, with repeated interviews the false stories quickly came to resemble the true stories in terms of the number of details mentioned, the spontaneity of the utterances, the number of new details, inconsistency across narratives, the elaborativeness of the details, and the cohesiveness of the narrative. It is only consistency across tellings that differentiates true from false narratives. A word of caution on this observation is in order, however: It seems likely that consistency across narratives could become a less potent predictor if children are repeatedly interviewed. In addition, when false stories are told as a result of repeated suggestive interviewing, they take on additional qualities that make them seem more believable than true narratives. Specifically, after a number of interviews, false narratives contained more descriptive material than did true narratives.

Finally, the present results indicate that fine-grained analyses of false narratives produce similar patterns of results for unpleasant (negative) events as for pleasant

(positive) events. Although some might argue that we obtained high rates of assents to false-negative events because the child was not a participant, it is not clear that our most recent data support this hypothesis. Although it was only suggested to the children that they might have seen a theft, many of those who did assent falsely created narratives in which they were participants. Children reported chasing the thief, being chased by the thief, hitting the thief, and similar types of actions. Therefore, it appears that if children are interviewed with a combination of suggestive techniques, many will assent to and create complex narratives of false-negative events.

Summary of Suggestibility Studies and Directions for New Research

Although we have concentrated on the conditions that can compromise reliable reporting, it is also important to acknowledge that under certain circumstances children are capable of providing accurate, detailed, and useful information. For example, in many of our own studies, children in the control group conditions (who were questioned in nonsuggestive interviews) often recalled events flawlessly (e.g., Bruck, Ceci, & Francoeur, 1995; Ceci, Loftus, et al., 1994; Leichtman & Ceci, 1995). This indicates that the absence of suggestive techniques allows even very young preschoolers to provide highly accurate reports, although they may be sparse in the number of details. Also, the results of a number of studies of children's autobiographical recall or memory for events indicate that children's recall is at times highly accurate and at times quite detailed about a large range of events (Baker-Ward, Gordon, Ornstein, Larus, & Clubb, 1993; Parker, Bahrick, Lundy, Fivush, & Levitt, in press; Peterson & Bell, 1996). What characterizes these studies is the neutral tone of the interviewer, the limited use of misleading questions (for the most part, if suggestions are used, they are limited to a single occasion), and the absence of the induction of any motive or bias for the child to make a false report.

An important implication of studies that emphasize the strength of children's memories is that they highlight the conditions under which children should be interviewed if one wishes to obtain reliable reports. Again, when children are interviewed by unbiased, neutral interviewers, when the number of interviews as well as the number of leading questions are kept to a minimum, and when there is an absence of threats, bribes, and peer pressure, then children's reports are at considerably less risk for taint. These are the conditions we must strive for when eliciting information from young children.

We also have not placed much emphasis on age differences in children's suggestibility. Many of the studies that we have reviewed in this article are not developmental in nature, in that only one age group, preschoolers, was included. However, there are other studies that do involve age group comparisons, and these overwhelmingly show that preschoolers are the most suggestible group. This conclusion is based on a previous literature review (Ceci & Bruck, 1993) that reported that in approximately 88% of the studies (14 out of 16) that involved comparisons of preschoolers to older children or to adults, preschool children were the most suggestible group. Since that publication, new studies on children's suggestibility are being published on a regular basis; these newer data continue the trend reported in 1993, with approximately four out of five studies demonstrating significant age differences in suggestibility.

Despite these strong age differences, it is nonetheless important to point out that it is premature and possibly even wrong to assume that only preschool children are suggestible or that there should be no concern about the reliability of older children's testimony if they are subjected to suggestive interviews. There are two prongs to this argument. First, developmental studies may underestimate the suggestibility of older children. That is, when preschoolers are included in the same study as older children, the task is usually designed to be suitable for the preschool children. As a result, the task often may be too easy for older participants, thus creating ceiling effects. It is difficult to create experimental situations that have the same meaning, interest, and difficulty for all age groups.

The second prong to the argument that older children are suggestible is that there is ample evidence that children do not reach adult levels of resistance to erroneous suggestions prior to early adolescence. There are some studies (see reviews by Ceci & Bruck, 1993, 1995) that include older children, and in these studies, the impact of suggestive interviewing techniques is frequently marked. For example, in some studies, 8-year-old, 9-year-old, and even 10-year-old children are significantly more suggestible than adults (Ackil & Zaragoza, 1995; Warren & Lane, 1995). Although it could be claimed that these studies have less relevance for the courtroom because they assessed children's memories of neutral events that did not involve participation, similar findings have been reported when children were suggestively interviewed about events in which they themselves participated. For example, when asked to recall the details of an event that occurred four years previously, children between the ages of 7 and 10 years were influenced by the atmosphere of accusation created by the experimenters, and they inaccurately reported events (Goodman, Wilson, Hazan, & Reed, 1989). In another study, a significant proportion of 8-year-olds reported that something "yuckie" was placed in their mouths when these nonevents were incorporated into stories that parents read to their children (Poole & Lindsay, 1996).

Even adults' recollections are impaired by suggestive interviewing techniques, albeit in reduced magnitude compared with the impairment obtained for very young children. For example, Loftus and Pickrell (1995) implanted false memories of being lost in a shopping mall in 25% of their participants. In other studies, Hyman and his colleagues implanted false childhood memories of spilling a punch bowl at a wedding for 38% of adult participants (Hyman & Pentland, 1996) or of being hospitalized for an ear infection in 20% of adult participants

(Hyman, Husband, & Billings, 1995). Malinoski and Lynn (1995, 1996) showed that one can easily age-regress normal college students to remember events from before the first year of life. Interestingly, the success of memory implantations in adult participants seems to hinge on many of the same factors that we found to be suggestive in our interviews with children. For example, repetition of the misinformation, visual imagery induction, and the status of the suggester (interviewer) are important factors.

On the basis of this brief review, we conclude that there are solid reasons to assume that preschoolers are not the only age group that can be influenced by suggestive interviewing techniques. Thus, it is also important to extend the research to develop newer paradigms for middle childhood and even adolescence to examine the magnitude, boundary conditions, and factors (as discussed below) involved in the suggestibility of individuals of these age groups, which have been as neglected today as the preschool age group was during the first half of this century.

It should not be surprising that children can be influenced to give inaccurate reports—a finding that many seem to find depressing, negative, and disconcerting. Although some have tried to dilute this finding by referring to the strengths of children's memory and reporting, it is equally true that there are pronounced developmental differences that contribute to reporting, beliefs, and memory. There are developmental differences in the degree to which children accurately encode, store, and retrieve memories (Brainerd & Ornstein, 1991). There are developmental differences in forgetting, retention, and relearning curves (Brainerd, Reyna, Howe, & Kingma, 1990). Young children are especially prone to making source misattributions (Ackil & Zaragoza, 1995; Parker, 1995; Poole & Lindsay, 1996). Each of these factors is thought to contribute to suggestibility.

In the noncognitive domain, there are also developmental differences in social compliance and in the willingness to please an authority figure and to provide information when requested for it. Generally, young children believe adults and accept their statements as credible (Ackerman, 1983; Sonnenschein & Whitehurst, 1980). Thus, young children are more suggestible when interviewed by an adult than by a peer (Ceci, Ross, & Toglia, 1987), and they are more likely to be swayed when interviewed by an adult of high prestige or authority (Tobey & Goodman, 1992).

Although such frameworks for understanding developmental differences in suggestibility abound, there has been little research to provide empirical support for many of the hypothesized relationships. This is not to say that there has been no research. For example, some researchers have attempted to determine the relative importance of social versus cognitive factors in accounting for suggestibility effects (e.g., Ceci et al., 1987; Zaragoza, Dahlgren, & Muench, 1992). The results of these studies are inconsistent, and the issue as to the ascendance of one factor over the other remains unresolved. As we have argued in previous work (Ceci & Bruck, 1993), it seems that for now we can conclude that although social factors (e.g., desire to please powerful authority figures) are quite important, they do not appear to fully account for all suggestibility effects, and they probably interact with cognitive factors (strength of the memory trace, knowledge representations) to boost suggestibility in certain circumstances. For example, when an authority source is perceived as more powerful, children may pay greater attention to the erroneous message, thus allowing more of it to seep into their subsequent recollection. Conversely, it is possible that the degree to which social factors play a role has a cognitive basis; when memory traces are weak or nonexistent, children may be more compliant and willing to accept suggestions because there is no competing memory trace to challenge the suggestion.

We expect that in the next decade, researchers will devote much effort to developing and testing theoretical models of the mechanisms underlying suggestibility. By virtue of the phenomenon itself, the ensuing models, by necessity, will be complex and multidimensional, involving not only a host of cognitive factors (trace strength, the nature of the mnemonic representation, strategy use) and social factors (interviewer status, pressures to comply) but also biological factors (physiological reactivity). Moreover, we anticipate that the best models will also account for individual differences in suggestibility. As such, they may include personality factors (e.g., compliance, field independence, self-esteem), task factors (e.g., whether the test is retrieval intensive, number of exposures), and even demographic factors.

Policy Implications

The major implication that we see from this field of research can be simply summarized in five words: "training, training, and more training." Although the results that we report in this article have been replicated in a number of different laboratories and although they are accepted by many social scientists, interestingly, these data have not filtered down to the very communities that might most profit from them—those professionals who interview, assess, and treat children. It is surprising how little training frontline professionals receive in terms of interviewing children. When they are given guidelines, these rarely include any advice about how to be cautious and how to avoid contamination. Thus, it is not sufficient that professionals be provided with training; the training programs themselves must contain the most up-to-date and relevant materials. In the past five years, we have talked to dozens of groups of professionals involved in interviewing children, such as child protective service workers, law enforcement officials, and therapists. It is always surprising after each of these presentations to be told by experienced, highly placed professionals that this is the first time they have ever heard about this type of research.

To illustrate, in the summer of 1996, 26 child protection workers from around the world came to Cornell University for a two-week training conference. The majority of the participants had received a master's degree or better, and more than half were in positions of training,

program development, or supervision, or were senior caseworkers. All but one of these individuals had five or more years of experience in the field. Before the training program, participants' practices and knowledge regarding the interviewing of children and of child development were assessed by means of a questionnaire.

Although these individuals reported an average of 75 hours of training in interviewing techniques, child development issues, or both, almost one third reported no knowledge of the scientific literature related to interviewing or to child development. Of the remaining individuals who claimed to possess some knowledge of the scientific literature, the proportion of questions that they answered correctly was generally very low. For example, most participants agreed with the statement "Interviewers who believe that abuse has occurred should reinterview children who initially deny the abuse." Only 8 individuals reported knowing the research in this area, and their responses were relevant only 25% of the time. The results for questions concerning the use of selective reinforcement were much better; 90% of participants stated that the following statement was problematic: "In an interview it is important to be supportive of a child suspected of abuse by providing such statements as 'Don't be afraid to tell me what she did to you,' 'It's O.K. to tell, you will feel better once you have told me.' " However, only 10 individuals reported knowing the scientific evidence on this issue. Similarly, more than 50% of the participants had no knowledge of the scientific literature on developmental changes in suggestibility and on the use of anatomical dolls, but even when participants did state knowledge of the scientific literature, only 60% of these accurately stated that younger children are most suggestible and that children's play with anatomical dolls is not diagnostic of sexual abuse. Finally, for statements regarding the boundary conditions for suggestibility (e.g., preschoolers are more suggestible than older children for central details and for peripheral details), although more individuals claimed to possess scientific knowledge in this more than any other area, again the accuracy rate was very low, just over 30%.

Making the research available to practitioners is only a very first step. Even when the research is made available, long-existent beliefs and biases continue to influence not only what individuals objectively state that they know but also what they practice. It is very difficult to translate research into effective training programs, and such programs have a history of falling short (Doris, Mazur, & Thomas, 1995). The existing data indicate that good interviewing skills cannot be taught effectively in a short period of time (e.g., a 10-day session), and often interviewers fall back on old, ineffectual (and potentially dangerous), albeit more comfortable, habits after a period of time (Memon, Bull, & Smith, 1995). Presumably, much more time is needed for rehearsal with feedback of the trained skills—a luxury not readily afforded by most investigative agencies. Regardless of the difficulty of meeting these objectives, until professionals become familiar with this field of research and until they institute effective training programs, then we should expect to face a steady stream of cases that hinge on the reliability of the child's testimony and on the competence of the child's interviewers.

The second major policy implication of the research on the reliability of children's statements is the need to develop scientifically validated interviews. The ideal interviews will have to meet a number of criteria. First, they need to incorporate techniques that have been shown to be beneficial in eliciting complete and accurate testimony from children. For example, Saywitz and Snyder (1996) have developed a procedure to expand children's spontaneous reports by teaching them a narrative elaboration procedure. Sternberg et al. (1996) developed a rapport-building procedure that eventuated in children providing more spontaneous information. Carter, Bottoms, and Levine (1996) have delineated some components of a supportive interviewing technique that result in children's increased resistance to misinformation. At the same time, the ideal interview should not contain techniques that have been found to have harmful consequences (e.g., the use of anatomically detailed dolls, the use of props, the use of guided imagery). Many of these potentially detrimental techniques have been described in this article.

Next, the ideal interview must integrate or select these various techniques into a logical manner and age-appropriate format. That is, although a number of different studies point to the beneficial effects of certain types of procedures, one must be sure that each of these procedures continues to be effective when combined with other procedures. Furthermore, according to some methodologists, the ideal protocol should also contain the minimal number of techniques and phases that are necessary to elicit the required information. This criterion is sometimes referred to as *incremental validity* (see Wolfner, Faust & Dawes, 1993, for the application of this reasoning to interviews of child witnesses). Thus, in its development phase, the protocol would have to be tested to ensure that the inclusion of each element within the interview significantly enhances the amount of information obtained. If, when compared with some other element, an element does not increase information or actually results in less information, then that element should not be included as part of the protocol. Although it may seem harmless to allow such elements to remain, to do so would not facilitate the protocol's practical application. Sometimes it is very important that an interview be completed as expeditiously as possible, which means that a scientifically validated protocol should contain exactly the best combination of pieces, no more or less. This procedure is critical not only to ensure construct validity but also because of the exigencies that often accompany child interviews.

At present, there are a number of different interview protocols at various stages of development and validation (see Poole & Lamb, in press, for a full description). These include the Cognitive Interview (Fisher & Geiselman, 1992), the Step-Wise Interview (Yuille, Hunter, Joffe, & Zaparniuk, 1993), the Structured Interview (Memon,

Cronin, Eaves, & Bull, 1993), and the National Institute of Child Health and Human Development protocol (Lamb, Sternberg, Esplin, Hershkowitz & Orbach, 1997). Each of these has been developed on scientific premises and is being tested in a variety of settings. One expects that these protocols will result in more systematic and improved interviews with young children.

Conclusions

At the beginning of this article, we discussed the reliability of young children's reports by focusing on some problems in the accurate diagnosis of child sexual abuse. Of particular concern is the pattern of disclosure in which the child initially denies but then later reports abuse. Although some professionals state that this is a common pattern among sexually abused child victims, the research reviewed in this article provides an alternative hypothesis that should be considered on a case-by-case basis: Professionals should be alerted to the possible contaminating effects of suggestive interviewing techniques. We maintain that our discussion is not specific to child sexual abuse but applies also to other situations in which the child witness replies with denials to initial questions and only makes a serious allegation with suggestive questioning.

Although we have shown that to some degree young children's false allegations that emerge as a result of suggestive interviewing practices reflect cognitive and social factors characteristic of young children, it is also the case that factors external to the children—those that characterize suggestive interviews—probably have a predominant influence on the emergence of children's false allegations. Thus, we do not view the young child as an incompetent witness. Rather, if questioned under the appropriate circumstances, the young witness may provide the court with forensically important evidence. Because we view external rather than internal psychological factors as most predictive of producing false allegations, we urge the development of appropriate interviewing schedules and training programs for professional interviewers.

REFERENCES

Ackerman, B. (1983). Speaker bias in children's evaluation of the external consistency of statements. *Journal of Experimental Child Psychology, 35,* 111–127.

Ackil, J. K., & Zaragoza, M. S. (1995). Developmental differences in eyewitness suggestibility and memory for source. *Journal of Experimental Child Psychology, 60,* 57–83.

American Humane Association. (1988). *Highlights of official child neglect and abuse reporting, 1986.* Denver, CO: Author.

Baker-Ward, L., Gordon, B., Ornstein, P. A., Larus, D., & Clubb, P. (1993). Young children's long-term retention of a pediatric examination. *Child Development, 64,* 1519–1533.

Bays, J., & Chadwick, C. (1993). Medical diagnosis of the sexually abused child. *Child Abuse & Neglect, 17,* 91–110.

Berenson, A., Heger, A., & Andrews, S. (1991). Appearance of the hymen in newborns. *Pediatrics, 87,* 458–465.

Bottoms, B., & Goodman, G. (Eds.). (1996). *International perspectives on child abuse and children's testimony.* Thousand Oaks, CA: Sage.

Bottoms, B., Shaver, P., & Goodman, G. (1996). An analysis of ritualistic and religion-related child abuse allegations. *Law & Human Behavior, 20,* 1–34.

Bradley, A., & Wood, J. (1996). How do children tell? The disclosure process in child sexual abuse. *Child Abuse & Neglect, 20,* 881–891.

Brainerd, C. J., & Ornstein, P. A. (1991). Children's memory for witnessed events: The developmental backdrop. In J. L. Doris (Ed.), *The suggestibility of children's recollections* (pp. 10–20). Washington, DC: American Psychological Association.

Brainerd, C., Reyna, V. F., Howe, M. L., & Kingma, J. (1990). The development of forgetting and reminiscence. *Monographs of the Society for Research in Child Development, 55*(3–4, Serial No. 222), 1–93.

Bruck, M., Ceci, S. J., & Francoeur, E. (1995, March). *Anatomically detailed dolls do not facilitate preschoolers' reports of touching.* Paper presented at the annual meeting of the Society for Research in Child Development, Indianapolis, IN.

Bruck, M., Ceci, S. J., Francoeur, E., & Barr, R. J. (1995). "I hardly cried when I got my shot!": Influencing children's reports about a visit to their pediatrician. *Child Development, 66,* 193–208.

Bruck, M., Ceci, S. J., Francoeur, E., & Renick, A. (1995). Anatomically detailed dolls do not facilitate preschoolers' reports of a pediatric examination involving genital touching. *Journal of Experimental Psychology: Applied, 1,* 95–109.

Bruck, M., Ceci, S. J., & Hembrooke, H. (1997). Children's reports of pleasant and unpleasant events. In D. Read and S. Lindsay (Eds.), *Recollections of trauma: Scientific research and clinical practice* (pp. 199–219). New York: Plenum Press.

Bull, R., & Cherryman, J. (1995). *Helping to identify skills gaps in specialist investigative interviewing: Enhancement of professional skills.* UK: Home Office Police Department.

Carter, C. A., Bottoms, B. L., & Levine, M. (1996). Linguistic and socioemotional influences on the accuracy of children's reports. *Law & Human Behavior, 20,* 335–358.

Ceci, S. J., & Bruck, M. (1993). The suggestibility of the child witness: A historical review and synthesis. *Psychological Bulletin, 113,* 403–439.

Ceci, S. J., & Bruck, M. (1995). *Jeopardy in the courtroom: A scientific analysis of children's testimony.* Washington, DC: American Psychological Association.

Ceci, S. J., Bruck, M., & Rosenthal, R. (1995). Children's allegations of sexual abuse: Forensic and scientific issues: A reply to commentators. *Psychology, Public Policy, and Law, 1,* 494–520.

Ceci, S. J., Crotteau-Huffman, M., Smith, E., & Loftus, E. W. (1994). Repeatedly thinking about non-events. *Consciousness & Cognition, 3,* 388–407.

Ceci, S. J., Loftus, E. W., Leichtman, M., & Bruck, M. (1994). The role of source misattributions in the creation of false beliefs among preschoolers. *International Journal of Clinical and Experimental Hypnosis, 62,* 304–320.

Ceci, S. J., Ross, D., & Toglia, M. (1987). Age differences in suggestibility: Psycholegal implications. *Journal of Experimental Psychology: General, 117,* 38–49.

Conerly, S. (1986). Assessment of suspected child abuse. In K. MacFarlane, J. Waterman, S. Conerly, L. Damon, M. Durfee, & S. Long (Eds.), *Sexual abuse of young children: Evaluation and treatment* (pp. 30–51). New York: Guilford Press.

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, 78,* 428–437.

Davies, G., Lloyd-Bostock, S., McMurran, M., & Wilson, C. (Eds.). (1995). *Psychology, law, and criminal justice: International developments in research and practice.* Berlin, Germany: Walter de Gruyter.

Doris, J., Mazur, R., & Thomas, M. (1995). Training in child protective services: A commentary on the amicus brief of Bruck and Ceci (1993/1995). *Psychology, Public Policy, and Law, 1,* 479–493.

Finlayson, L. M., & Koocher, G. P. (1991). Professional judgment and child abuse reporting in sexual abuse cases. *Professional Psychology: Research and Practice, 22,* 464–472.

Fisher, R. P., & Geiselman, R. E. (1992). *Memory-enhancing techniques for investigative interviewing: The cognitive interview.* Springfield, IL: Charles C Thomas.

Fivush, R., Haden, C., & Adam, S. (1995). Structure and coherence of preschoolers' personal narratives over time: Implications for childhood amnesia. *Journal of Experimental Child Psychology, 60,* 32–56.

Gonzalez, L., Waterman, J., Kelly, R., McCord, J., & Oliveri, M. (1993). Children's patterns of disclosures and recantations of sexual and ritualistic abuse allegations in psychotherapy. *Child Abuse & Neglect, 17,* 281–289.

Goodman, G. S., Taub, E. P., Jones, D. P., England, P., Port, L., Rudy, L., & Prado, L. (1992). Testifying in criminal court. *Monographs of the Society for Research in Child Development, 57*(5, Serial No. 229), 1–142.

Goodman, G. S., Wilson, M. E., Hazan, C., & Reed, R. S. (1989, April). *Children's testimony nearly four years after an event.* Paper presented at the annual meeting of the Eastern Psychological Association, Boston, MA.

Horner, T. M., Guyer, M. J., & Kalter, N. M. (1993a). The biases of child sexual abuse experts: Believing is seeing. *Bulletin of the American Academy of Psychiatry and Law, 21,* 281–292.

Horner, T. M., Guyer, M. J., & Kalter, N. M. (1993b). Clinical expertise and the assessment of child sexual abuse. *Journal of the American Academy of Child and Adolescent Psychiatry, 32,* 925–931.

Hulse, D. A. (1994). *Linguistic complexity in child abuse interviews.* Unpublished master's thesis, University of Tennessee at Chattanooga.

Hyman, I., Husband, T., & Billings, F. (1995). False memories of childhood experiences. *Applied Cognitive Psychology, 9,* 181–197.

Hyman, I., & Pentland, J. (1996). The role of mental imagery in the creation of false childhood memories. *Journal of Memory & Language, 35,* 101–117.

Idaho v. Wright, 497 U.S. 805 (1990).

Jones, D., & McGraw, J. M. (1987). Reliable and fictitious accounts of sexual abuse in children. *Journal of Interpersonal Violence, 2,* 27–45.

Kalichman, S. C. (1993). *Mandated reporting of suspected child abuse: Ethics, law, and policy.* Washington, DC: American Psychological Association.

Kendall-Tackett, K. A., Williams, L. M., & Finkelhor, D. (1993). Impact of sexual abuse on children: A review and synthesis of recent empirical studies. *Psychological Bulletin, 113,* 164–180.

La Fontaine, J. S. (1994). *The extent and nature of organised and ritual abuse. Research findings.* Department of Health. London: HMSO.

Lamb, M. E., Hershkowitz, I., Sternberg, K. J., Esplin, P. W., Hovav, M., Manor, T., & Yudilevitch, L. (1996). Effects of investigative utterance types on Israeli children's responses. *International Journal of Behavioral Development, 19,* 627–637.

Lamb, M. E., Sternberg, K. J., Esplin, P. W., Hershkowitz, I., & Orbach, Y. (1997). Assessing the credibility of children's allegations of sexual abuse: A survey of recent research. *Learning & Individual Differences, 9,* 175–194.

Leichtman, M. D., & Ceci, S. J. (1995). The effects of stereotypes and suggestions on preschoolers' reports. *Developmental Psychology, 31,* 568–578.

Leippe, M., Romanczyk, A., & Manion, A. P. (1991). Eyewitness memory for a touching experience: Accuracy and communication style differences between child and adult witnesses. *Journal of Applied Psychology, 76,* 367–379.

Lipmann, O., & Wendriner, E. (1906). Aussageexperimente im Kindergarten. *Beitr. zur Psycologie der Aussage*(Stern), *2,* S418–S423.

Loftus, E. F., & Pickrell, J. (1995). The formation of false memories. *Psychiatric Annals, 25,* 720–725.

MacFarlane, K., & Krebs, S. (1986). Techniques for interviewing and evidence gathering. In K. MacFarlane, J. Waterman, S. Conerly, L. Damon, M. Durfee, & S. Long (Eds.), *Sexual abuse of young children: Evaluation and treatment* (pp. 67–100). New York: Guilford Press.

Malinoski, P., & Lynn, S. J. (1995, August). *The pliability of early memory reports.* Paper presented at the 103rd Annual Convention of the American Psychological Association, New York, NY.

Malinoski, P., & Lynn, S. J. (1996, November). *The temporal stability of early memory reports.* Paper presented at the annual convention of the Society for Clinical and Experimental Hypnosis, Tampa, FL.

Maryland v. Craig, 110 S. Ct. 3157 (1990).

McGough, L. (1994). *Fragile voices: The child witness in American courts.* New Haven, CT: Yale University Press.

Melton, G. (1992). Children as partners for justice: Next steps for developmentalists. *Monographs of the Society for Research in Child Development, 57*(5, Serial No. 229), 153–159.

Melton, G., & Thompson, R. (1987). Getting out of a rut: Detours to less traveled paths in child-witness research. In S. J. Ceci, M. P. Toglia, & D. F. Ross (Eds.), *Children's eyewitness memory* (pp. 209–229). New York: Springer-Verlag.

Memon, A., Bull, R., & Smith, M. (1995). Improving the quality of the police interview: Can training in the use of cognitive techniques help? *Policy and Society, 5,* 53–68.

Memon, A., Cronin, O., Eaves, R., & Bull, R. (1993). The cognitive interview and child witnesses. *Issues in Criminological & Legal Psychology, 20,* 3–9.

Nathan, D., & Snedeker, M. (1995). *Satan's silence: Ritual abuse and the making of a modern American witch hunt.* New York: Basic Books.

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.

Parker, J. (1995). Age differences in source monitoring of performed and imagined actions on immediate and delayed tests. *Journal of Experimental Child Psychology, 60,* 84–101.

Parker, J., Bahrick, L., Lundy, B., Fivush, R., & Levitt, M. (in press). Effects of stress on children's memory for a natural disaster. In C. P. Thompson, D. J. Herrmann, J. D. Read, D. Bruce, D. G. Payne, & M. P. Toglia (Eds.), *Eyewitness memory: Theoretical and applied perspectives.* Mahwah, NJ: Erlbaum.

Peterson, C., & Bell, M. (1996). Children's memory for traumatic injury. *Child Development, 67,* 3045–3070.

Poole, D. A., & Lamb, M. E. (in press). *Investigative interviews of children: A guide for helping professionals.* Washington, DC: American Psychological Association.

Poole, D. A., & Lindsay, D. S. (1996, June). *Effects of parental suggestions, interviewing techniques, and age on young children's event reports.* Paper presented at the NATO Advanced Study Institute, Recollections of trauma: Scientific research and clinical practice, Port de Bourgenay, France.

Poole, D. A., & Lindsay, D. S. (in press). Assessing the accuracy of young children's reports: Lessons from the investigation of child sexual abuse. *Journal of Applied and Preventive Psychology.*

Rawls, J. (1996). How question form and body-parts diagrams can affect the content of young children's disclosures. Paper presented at the NATO Advanced Study Institute, Recollections of trauma: Scientific research and clinical practice, Port de Bourgenay, France.

Salmon, K., & Pipe, M. E. (1997). Providing props to facilitate young children's event recall: The impact of a one year delay. *Journal of Experimental Child Psychology, 65,* 261–292.

Saywitz, K., Goodman, G., Nicholas, G., & Moan, S. (1991). Children's memory of a physical examination involving genital touch: Implications for reports of child sexual abuse. *Journal of Consulting and Clinical Psychology, 59,* 682–691.

Saywitz, K. J., & Snyder, L. (1996). Narrative elaboration: Test of a new procedure for interviewing children. *Journal of Consulting and Clinical Psychology, 64,* 1347–1357.

Sedlak, A., & Broadhurst, D. (1996). *Executive summary of the Third National Incidence Study of Child Abuse and Neglect.* Washington, DC: U.S. Department of Health and Human Services.

Sonnenschein, S., & Whitehurst, G. (1980). The development of communication: When a bad model makes a good teacher. *Journal of Experimental Child Psychology, 3,* 371–390.

Sorensen, T., & Snow, B. (1991). How children tell: The process of disclosure of child sexual abuse. *Child Welfare, 70,* 3–15.

State v. Michaels, 136 N.J. 299, 642 A.2d 1372 (N.J., 1994).

Sternberg, K. J., Lamb, M. E., Hershkowitz, I., Yudilevitch, L., Orback, Y., Esplin, P., & Horav, M. (1996). Effects of introductory style on children's abilities to describe experiences of sexual abuse. *International Journal of Behavioral Development, 19,* 627–637.

Steward, M. S., & Steward, D. S., with Farquahar, L., Myers, J. E. B., Reinart, M., Welker, J., Joye, N., Driskll, J., & Morgan, J. (1996). Interviewing young children about body touch and handling. *Monographs of the Society for Research in Child Development, 61*(4–5, Serial No. 248).

Summit, R. (1983). The child sexual abuse accommodation syndrome. *Child Abuse & Neglect, 7,* 177–193.

Tobey, A., & Goodman, G. S. (1992). Children's eyewitness memory: Effects of participation and forensic context. *Child Abuse & Neglect, 16,* 779–796.

United States v. Rouse, F.3d 360 (8th Cir. 1996).

U.S. Department of Health and Human Services, National Center on Child Abuse and Neglect. (1996). *Child maltreatment 1994: Reports from the states to the National Center on Child Abuse and Neglect.* Washington, DC: U.S. Government Printing Office.

U.S. Department of Health and Human Services, National Center on Child Abuse and Neglect. (1997). *Child maltreatment 1995: Reports from the states to the National Center on Child Abuse and Neglect.* Washington, DC: U.S. Government Printing Office.

Wang, C. T., & Daro, D. (1996). *Current trends in child abuse reporting and fatalities: The results of the 1996 Annual Fifty State Survey.* Chicago: National Committee to Prevent Child Abuse.

Warren, A. R., & Lane, P. (1995). The effects of timing and type of questioning on eyewitness accuracy and suggestibility. In M. Zargoza (Ed.), *Memory and testimony in the child witness* (pp. 44–60). Thousand Oaks, CA: Sage.

Warren, A. R., Woodall, C. E., Hunt, J. S., & Perry, N. W. (1996). "It sounds good in theory, but . . .", Do investigative interviewers follow guidelines based on memory research? *Child Maltreatment, 1,* 231–245.

White v. Illinois, 502 U.S., 112 S. Ct. 736 (1992).

Wolfner, G., Faust, D., & Dawes, R. (1993). The use of anatomical dolls in sexual abuse evaluations: The state of the science. *Applied and Preventative Psychology, 2,* 1–11.

Yuille, J. C., Hunter, R., Joffe, R., & Zaparniuk, J. (1993). Interviewing children in sexual abuse cases. In G. S. Goodman & B. Bottoms (Eds.), *Child victims, child witnesses* (pp. 95–116). New York: Guilford Press.

Yuille, J., Marxsen, D., & Menard, K. (1993). *Interviewing and assessing children in sexual abuse investigations: A field study.* Unpublished report for the Ministry of Social Services, Victoria, British Columbia, Canada.

Zaragoza, M., Dahlgren, D., & Muench, J. (1992). The role of memory impairment in children's suggestibility. In M. L. Howe, C. J. Brainerd, & V. F. Reyna (Eds.), *Development of long-term retention* (pp. 184–216). New York: Springer-Verlag.