Reliability of Child Witnesses' Reports

Maggie Bruck and Stephen J. Ceci

When children report being a victim of, or a witness to, a crime, two primary sets of issues arise in evaluating their testimony. The first set of issues concerns the memories of children who have been actual witnesses to, or victims of, a trauma. Topics related to this set of issues include the cognitive, motivational, and emotional factors that influence the accuracy of the child's report of the traumatic event in question. The research in this area is based on the assumption that the child has actually experienced the traumatic event.

The second set of issues, and the focus of the present chapter, concerns whether it is possible to elicit false reports from a child about nonexperienced traumatic events; that is, events that would be traumatic had the child been a witness to, or victim of, them. The major topics in this area of research include the conditions that precipitate false reports (e.g., suggestive interviews), the psychological status of false reports (false beliefs vs. lies), and developmental trends in both. In this chapter we focus on forensic techniques that may elicit false reports.

CORE ISSUES AND CONTROVERSIES

Distinguishing Suggestive and Neutral Interviews: Misleading and Open-Ended Questions

The Nature of the Controversy

As a result of more than a decade of intensive study on the factors that affect the reliability of children's testimony, it is fairly well agreed by all professionals that suggestive techniques can compromise the accuracy of children's reports. For most professionals (and laypersons), suggestive techniques are synonymous with the use of misleading questions, whereas nonsuggestive techniques are equated with open-ended questions to obtain information from the child.

The controversy concerns whether the absence of misleading questions and the sole use of open-ended questions constitute "a safe interview." We have argued (see Bruck, Ceci, & Principe, 2006; Ceci & Bruck, 1995) that the number of leading versus open-ended questions is not a good index of the suggestiveness of an interview; rather, the best interviews are characterized by the absence of (1) what we have termed "interviewer bias" and (2) a number of specific suggestive techniques.

Science Relevant to the Controversy: Characteristics of a Suggestive Interview

According to Bruck and Ceci's model (Bruck et al., 2006), interviewer bias is the central characteristic that drives the structure of suggestive interviews. Interviewer bias characterizes interviewers who hold *a priori* beliefs about the occurrence of certain events and, as a result, conduct their interviews so as to obtain confirmatory evidence for these beliefs without considering plausible alternative hypotheses. When children provide such interviewers with inconsistent or bizarre evidence, it is either ignored or interpreted within the framework of the biased interviewer's initial hypothesis.

According to our model, interviewer bias influences the entire architecture of an interview and is revealed through a variety of suggestive interviewing techniques. One of the main dimensions of such interviews is the absence of open-ended questions ("Tell me about it; then what happened?"); thus, one does not hear the child's narrative but rather the child's monosyllabic responses to specific questions (e.g., "Did he ever touch you here?"; "Did he do it upstairs?"; "Was it upstairs or downstairs?"). The repetition of these questions (or themes of these questions) within and between inter-

views is quite suggestive; the repetition provides the child clues about the interviewer's beliefs.

Other suggestive techniques include repeated interviews (especially when the child does not provide satisfactory statements in the first interview), implicit or explicit threats, bribes, and rewards for the desired answer, stereotype induction (e.g., telling children the suspected perpetrator "does bad things"), the use of peer pressure (telling the child that other children have told the truth and now it is his or her turn), and guided imagery (asking children to create a mental picture of a specific event and to think about its details; see Bruck et al., 2006; Ceci & Bruck, 1995, for details). The use of nonverbal props, including anatomically detailed dolls, can also be suggestive, especially with preschool-age children (e.g., Bruck, Ceci, & Francoeur, 2000). Although each suggestive technique is associated with error, the risk for false statements is greatly augmented when interviews contain a combination of suggestive techniques that increases the salience of the interviewer's bias.

Below we provide two examples of scientific studies of the influence of interviewer bias on children's report accuracy, followed by a summary of some of the research on the effects of suggestive techniques on children's reports.

CHESTER THE JANITOR

Thompson, Clarke-Stewart, and Lepore (1997) conducted a study in which children ages 5–6 viewed a staged event that could be construed as either a misdeed or an innocent act. Some children interacted with a confederate named "Chester" as he cleaned some dolls and other toys in a playroom. Other children interacted with Chester as he handled the dolls roughly and in a mildly abusive manner. The children were then questioned about this event. The interviewer was either (1) "accusatory" (suggesting that the janitor had been inappropriately playing with the toys instead of working), (2) "exculpatory" (suggesting that the janitor was just cleaning the toys and not playing), or (3) "neutral" and nonsuggestive. Following the first interview, all children were asked to tell in their own words what they had witnessed and then they were asked questions about the event. Immediately after the interview and 2 weeks later, parents asked their children to recount what the janitor had done.

When questioned by a neutral interviewer or by an interviewer whose interpretation was consistent with the activity viewed by the child, children's accounts were both factually correct and consistent with the janitor's script.

However, when the interviewer was biased in a direction that contradicted the activity viewed by the children, those children's stories conformed to the suggestions or beliefs of the interviewer. In addition, children's answers to interpretive questions (e.g., "Was he doing his job or just being bad?") were in agreement with the interviewer's point of view, as opposed to what actually happened. When asked neutral questions by their parents, the children's answers remained consistent with the interviewers' biases.

SURPRISE PARTY

Bruck, Ceci, Melnyk, and Finkelberg (1999) showed how interviewer bias can quickly develop in natural interviewing situations, and how it not only taints the responses of child interviewees but also the reports of the adult interviewers. In their study, a special event was staged for 90 preschool children in their school. In groups of three children, and with the guidance of research assistant A, the children surprised research assistant B with a birthday party, played games, ate food, and watched magic tricks. Another 30 children did not attend the birthday party but in groups of two, they simply colored a picture with research assistants A and B. These children were told that it was assistant A's birthday and saw one of the magic tricks.

Interviewers (who were recruited from graduate degree programs in social work or counseling and who had training and experience in interviewing children) were asked to question four children individually about what had happened when special visitors came to the school. The interviewers were not told about the events but were simply told to find out from each child what had happened. The first three children that each interviewer questioned attended the birthday party and the fourth child attended the coloring event.

Bruck and colleagues (2006) found that the fourth child (who attended the coloring event and was interviewed last) produced twice as many errors as the children who attended the birthday party. For example, 60% of the children who only colored made false claims that involved attending a birthday party. This result suggests that the interviewers had built up a bias that all the children had attended a birthday party as a result of interviewing three consecutive children who actually had done so. By the time they interviewed the fourth child, they structured their interviews to elicit claims consistent with this hypothesis. Thus, if interviewers have the belief that all the children had experienced a certain event, then it is probable that many of the children will come to make such claims even though they were nonparticipants (or nonvictims). Another important finding from this study

was that even when the fourth child denied attending a birthday party, 84% of their interviewers later reported that all four of the children they interviewed had reported to them that they attended a birthday party. These data suggest that regardless of what children actually say, biased interviewers inaccurately report children's claims, rendering them consistent with their own hypotheses. Biased interviewers often do not have a conscious desire or goal to manipulate the testimony of their interviewee; they are unaware of the strategies they use.

These two studies and others like them (Bruck et al., 2006) provide evidence that interviewers' beliefs about an event can influence their judgments as well as their style of questioning. These beliefs, in turn, can affect the accuracy of children's testimony. Finally, interviewers' beliefs can affect their own ability to accurately recollect what children actually told them during the interview. These findings highlight the dangers of having only one hypothesis about the event in question, especially when this hypothesis is incorrect.

A review of actual forensic interviews with children reveals a number of interviewing techniques that may allow the child to infer the beliefs of the interviewer. Using a variety of techniques in the laboratory, children are interviewed about events they have and have not experienced. The accuracy of their reports is examined as a function of the types of techniques used in the interview. Briefly, when children are questioned about events they did not experience (e.g., seeing a thief steal food from the day care center; Bruck, Ceci, & Hembrooke, 2002) or about non-occurring details within experienced events (e.g., "The man put something yucky in your mouth"; Poole & Lindsay, 2001), their reports are more error-prone if suggestive techniques are used than if the questioning takes place in a neutral, nonsuggestive manner (see Bruck et al., 2006, for details).

Some studies show that interviews can be completely devoid of leading questions and still be quite suggestive because of the presence of interviewer bias. Conversely, sometimes leading questions may not pose a risk to the reliability of children's reports in the absence of interviewer bias. For example, a study by Garven, Wood, and Malpass (2000) illustrates both of these points. They asked kindergarten children to recall details when a visitor named Paco came to their classroom and read a story, gave out treats, and wore a funny hat. The children were asked misleading questions about plausible events (e.g., "Did Paco break a toy?") and about bizarre events (e.g., "Did Paco take you in a helicopter to a farm?"). Some of the children were given selective feedback after their answers to the misleading questions.

"No" responses were met with negative appraisals by the interviewer, as in the following exchange:

> INTERVIEWER: Did Paco take you somewhere in a helicopter?
>
> CHILD: No.
>
> INTERVIEWER: You're not doing good.

"Yes" responses were positively evaluated, as the following example illustrates:

> INTERVIEWER: Did Paco break a toy?
>
> CHILD: Yes.
>
> INTERVIEWER: Great; you're doing excellent now.

Children who were asked leading questions with selective reinforcement provided the desired but false answer to 35% of the plausible questions and to 52% of the bizarre questions. In contrast, a second group of children who did not receive this selective feedback falsely agreed with 13% of the plausible and 5% of the bizarre questions. Two weeks later, when the children were asked nonleading questions with no selective feedback, the same level of between-group differences was obtained. Thus, interviewer bias in a prior interview in the form of selective feedback had long-lasting negative effects on accuracy in a later unbiased interview.

There are two important points to bear in mind when analyzing transcripts of an interview: First, just because that particular interview may be neutral, prior interviews may have been suggestive, seeding false claims made in the second, neutral, interview. Second, the number of leading or suggestive questions deployed in an interview is not a good index of suggestiveness.

The Use of Suggestive Interviewing under Special Circumstances

The Nature of the Controversy: Child Sexual Abuse Accommodation Syndrome

There is a common belief among professionals who assess and treat children suspected of sexual abuse that it is necessary to engage in suggestive interviews because sexually abused children do not readily or spontaneously dis-

close abuse. The most popular embodiment of this idea is Summit's (1983) "child sexual abuse accommodation syndrome" (CSAAS). Roland Summit proposed that because of the traumatic characteristics of abuse, shame, and guilt, sexually abused children often delay or fail to disclose during child-hood and deny abuse even when asked. Furthermore, it is thought that when children do disclose, it is a slow process whereby they provide a few details and then recant these earlier claims.

Because of these circumstances, it is argued, it is necessary to use as many techniques as possible to extract disclosures in order to protect these otherwise silent children. Young children present the most concerns because, under most circumstances, they do not provide much information in response to open-ended questions; therefore, it is necessary to use a variety of strategies to elicit relevant information from this age group.

In contrast, some experts and professionals state that there is no scientific evidence to support the strong version of CSAAS, which is the foundation of the rationale for suggestive interviews, especially with young children. According to this position, it is important to develop structured nonsuggestive interviews that are feasible and effective for children of all ages.

Science Relevant to the Controversy: Empirical Support for the CSAAS Model

Because the CSAAS model was based on clinical intuitions rather than on data, we have reviewed the literature to determine its empirical support (London, Brück, Ceci, & Shuman, 2005; London, Bruck, Wright, & Ceci, 2008). We identified studies in which adults with histories of childhood abuse were asked to recall their disclosures in childhood. Across studies, an average of only 33% of the adults remembered disclosing the abuse in a timely fashion. These data support the CSAAS model insofar that sexually abused children are silent about their victimization and delay disclosure for long periods of time, sometimes even decades.

Although these studies are informative on the issue of delay of reporting, they are not informative on the issue of denial of abuse when asked. This is because the adult participants were never asked, "As a child, did anyone ever ask you or question you about abuse?" Thus, the data are silent on the phenomena of denial and recantation.

To address this aspect of the CSAAS model, we located another set of studies that provide some data relevant to this point. In the most recent article (London et al., 2008), we identified 18 studies conducted since 1990 that examined rates of denial and recantation by sexually abused children

who had been asked directly about abuse when they were assessed or treated at clinics for sexually abused children. The rates of denial during these assessment interviews were highly variable (4–76%), as were the rates of recantation (4–27%). We found that the methodological adequacy of each study was directly related to the denial and recantation rates observed: The weakest studies—those that included children who had made false allegations of abuse—produced the highest rates of denial, an average value of 61%. For example, the Sorenson and Snow (1991) study had the highest citation rate of all 18 studies, although it is the most methodologically compromised. Studies that included unrepresentative samples of abused children (e.g., those who had previously denied abuse) produced an average rate of denial of 51%. The set of studies that did not differentiate unfounded from founded cases produced a lower rate of denial, 31%. Finally, for the six methodologically superior studies, the average rate of denial was only 14%. The rate of recantation was also low for this last set of studies, 7%. These six studies provided denial and recantation data on children whose abuse status was considered "highly probable" (i.e., valid) and who were not selected because of special characteristics (e.g., sexually transmitted diseases [STDs]; peer abuse; single-parent families). These latter studies of sexually abused children's response patterns indicate that if they are directly asked, they do not deny their abuse, but disclose it. These findings lend no support to the notion that children who deny having been abused must be pursued with relentless suggestive questioning because otherwise they will not disclose the details of their abuse.

Although this analysis and conclusions about the empirical status of CSAAS have been generally accepted in the scientific community, they are not without detractors. Specifically, Lyon (2007) has claimed that the methods used to ensure valid cases of sexual abuse result in biased samples that exclude actual cases of abuse. To remedy the problem of valid diagnoses, Lyon reviewed the literature of the disclosure patterns of children with STDs. The average rate of denial across 21 studies was 57%. However, the generalizability of these findings are highly limited because only a small minority (2–3%) of sexually abused children have STDs and those who do have STDs are demographically very different from those who do not (see London et al., 2008). Also, most of the studies that Lyon cited in his survey were conducted before 1990 (when many of the current issues were not taken into account and thus resulted in less sophisticated studies), and most were not designed to examine disclosure patterns but rather the characteristics of children with STDs. Consequently, it is difficult to estimate disclosure rates from such data.

Lyon and colleagues (Malloy, Lyon, & Quas, 2007) also argued that recantation rates are much higher than calculated by London and colleagues (2005). To reach this conclusion they examined children (mainly Hispanic) facing dependency court hearings (i.e., removal from the home) as a consequence of their disclosures of abuse; 23% of these children recanted. Although this is one of the higher estimates of recantation, in actuality one might have predicted even higher rates given the dilemma of this special sample of children who were facing removal from the home and possible deportation or other punishment of a parent, unless they recanted. In addition, such children are often pressured by the nonoffending parent to withdraw their accusations. Thus even when there are a number of pressures to retract true allegations, only a minority of sexual abused children will do so. When these pressures are absent, recantation is rare.

Although most children will disclose abuse when directly asked, the question then becomes how to "directly ask" children. Because young children will provide more detailed answers to specific or cued questions than to open-ended questions, and because children's free recall is often sparse (Goodman & Reed, 1986; Peterson & Bell, 1996), a common view is that it is necessary to ask younger children specific and sometimes leading questions to elicit important information. With the acquisition of cognitive structures that organize events into coherent narratives, the need for specific questions declines with age (Kulkofsky, Wang, & Ceci, 2008). In addition, in the context of disclosures of sexual abuse, there is the belief that the child must be helped (through a variety of techniques) to overcome motivational and emotional barriers.

In fact, Bruck and colleagues (2002) found that highly suggestive interviews increased the chances that children would report otherwise embarrassing incidences of being punished. In this study, children were asked to tell an interviewer about a specific punishment they had recently experienced. The classroom teacher provided details of a punishment for each child (e.g., being placed in "time out" for calling other children bad names). When first asked in a neutral manner if the punishment had ever happened to them, 37% of the children answered negatively. A week later when suggestively interviewed (this interview involved the use of peer pressure, leading questions, repeated questions, positive feedback), 44% children still denied the punishment. But in the next interview, a week later, only 25% of the children denied, and by the next suggestive interview all children told about the embarrassing event. Thus, suggestive interviews promote children's reports about true events. However, these same techniques had the same effect on the elicitation of children's reports about never-experienced events. For

example, in this same study after two highly suggestive interviews, children not only assented to, but gave highly elaborate false reports about, helping a lady find her monkey in the park (only 6% of the children denied) and about witnessing a thief steal food from the day care (only 12% denied). Thus there are benefits but also very high risks to using suggestive interviewing techniques when one does not know what the child experienced; using such techniques, the interviewer will be successful in eliciting disclosures, but will not know whether these are accurate or inaccurate. Thus, the scientific literature indicates that under all circumstances (unless the truth is known, in which case there is no need to interview), unbiased, nonsuggestive interviews should be conducted.

We should note that, although a significant proportion of children denied a true event in this study, this figure should not be generalized to the proportion of sexually abused children who might deny abuse when first asked. This is because the motives to deny an actual punishment are quite different from denying sexual abuse. The former involve protecting oneself from revealing an embarrassing wrongdoing.

Research by Michael Lamb, Kathleen Sternberg, and colleagues has shown that the use of specific/leading questions is not necessary to elicit informative reports from children who are questioned about abuse (e.g., Lamb, Sternberg, & Esplin, 2000; Lamb et al., 2003; Sternberg et al., 1996). This team created a structured interview protocol and then trained interviewers of suspected child abuse victims in its use. The major feature of the protocol is to encourage the child to provide detailed life-event narratives through the guidance of open-ended questions (e.g., "Tell me what happened"; "And then what happened next?"; "And you said it happened at the store—tell me about the store"). The use of specific questions is allowed only after exhaustive free-recall prompting. Suggestive questions are highly discouraged. In one study, Lamb and colleagues (2003) examined the interviews of 16 trained police officers with 130 children (4–8 years old), all of whom had made allegations of sexual abuse. They found that 78% of preschoolers' allegations and disclosures were elicited through free-recall questions, and 66% of all children identified the suspect through open-ended questions (60% for preschoolers). These data dispel the belief that young children need to be bombarded with specific (suggestive) questions to elicit details of their traumatic events; in fact, children can provide detailed information through open-ended prompts, and, if a child denies abuse when asked directly, there is no scientifically compelling evidence that the child is "in denial." Abused children usually disclose when directly asked.

MYTHS AND MISCONCEPTIONS

Multiple Suggestive Interviews Are Needed to Taint a Report

To justify a suggestive interview, some experts have claimed that one suggestive interview is insufficient to taint a child's report and that taint only occurs when multiple suggestive techniques are used in repeated interviews (e.g., Ceci & Bruck, 1995).

This belief is not supported by the scientific literature, which shows that children can incorporate suggestions about salient events after a single interview (e.g., Bruck, London, Landa, & Goodman, 2007; Garven et al., 2000; Thompson et al., 1997). In the Paco study by Garven and colleagues (2000), for example, children's reports were significantly tainted after a single short suggestive interview. Other studies have directly compared the effects of multiple versus one suggestive interview on children's suggestibility. Evidence suggests that under a number of circumstances, one suggestive interview produces the same amount of taint as two or more suggestive interviews. The impact of a second interview depends on the spacing of the interviews from the initial events and from the final interview, and also on the strength of the original memory trace (Marche, 1999; Melnyk & Bruck, 2004; Powell & Thompson, 1997).

It is also true that children's reports can be tainted in the absence of highly suggestive, coercive techniques. Poole and Lindsay (2001) had parents read their children short narratives that outlined the children's previous encounters with a character known as Mr. Science at the researchers' laboratory. Unbeknownst to the parents, some of the details in the stories were inaccurate and thus were not experienced by the children when they met Mr. Science. Nonetheless, even under these mildly suggestive conditions, significant numbers of children (4- to 8-year-olds) later told an interviewer that they had experienced the suggested events (e.g., "The man put something yucky in my mouth").

Suggestibility Is Primarily a Problem for Preschoolers

Although much of the literature pays lip service to the concept that suggestibility exists at all ages, including in adults, the common view is that preschool children are disproportionately suggestible, and that there should be less concern about the tainting effects of suggestive interviews with older school-age children. This belief is reflected in the following expert testimony:

Well, in virtually all these studies, two- and three-year-olds do not do well in suggestibility, and the four- and five-year-olds do pretty well. It's true that the sorts of questioning that were asked of the children [in the case at hand] are not supported by basic research into suggestibility, but these children [in the case at hand] were all over the age of 6, the cut-off for suggestibility proneness in scientific studies. (expert testimony by a prosecution witness *In the matter of Riley Blanchard, Shelby Blanchard, and Austin Blanchard*, CA 2001; Tr. p. 1,441)

Yet although a number of findings show that elementary school-age children are less susceptible to suggestion than preschool children (e.g., Bright-Paul, Jarrold, & Wright, 2005; Ceci & Bruck, 1993; Ceci, Ross, & Toglia, 1987; Chae & Ceci, 2005; Holliday, 2003; Poole & Lindsay, 2001; Strange, Sutherland, & Garry, 2006), in most studies it is a matter of degree. That is, elementary school-age children show significant suggestibility effects even when preschoolers exhibit more suggestibility. Second, the results of a number of other studies show that not only is susceptibility to suggestion common in middle childhood, but that under some conditions there are small to no developmental differences in suggestibility. For example, Finnila and colleagues staged an event (a version of the Paco visit we described earlier) for 4- to 5-year-olds and 7- to 8-year-olds. One week later, half the children were given a low-pressure interview that contained some misleading questions with abuse themes (e.g., "He took your clothes off, didn't he?"). The other children received a high-pressure interview; they were told that their friends had answered the leading questions affirmatively, they were praised for assenting to the misleading questions, and when they did not assent, the question was repeated. In both conditions, there were no significant age differences in the percentage of misleading questions answered affirmatively, although a significant number (68%) were assented to in the high-pressure condition (Finnila, Mahlberga, Santtilaa, Sandnabbaa, & Niemib, 2003; see also Bruck et al., 2007).

Finally, under some conditions, older children are actually more suggestible than younger children (e.g., Ceci, Papierno, & Kulkofsky, 2007; Connolly & Price, 2006; Finnila et al., 2003; Schreiber & Parker, 2004). These reverse trends often reflect age differences in children's knowledge or what has been called "gist processing" (Reyna & Brainerd, 1998), such that higher levels of knowledge are sometimes associated with larger suggestibility effects. Sometimes, these reverse age trends are associated with social factors involving older children's greater sensitivity to the intent of some suggestive devices (e.g., a repeated question is a signal that the original answer

was incorrect or that the interviewer did not like the first answer). Brainerd, Reyna, and Ceci (2008) extensively reviewed the evidence for reverse developmental effects in which older children's greater knowledge renders them more vulnerable to suggestions than younger children. The bottom line is that, all age groups are vulnerable to misleading suggestions, even if preschoolers are sometimes disproportionately more vulnerable.

Children's Spontaneous Reports Are Always Accurate

A commonly held belief is that although children's prompted statements may be suspect, their spontaneous statements are generally accurate, and errors only occur in suggestive interviews. This belief has led some expert witnesses to opine that a child's statement must be correct because it was produced in response to a free-recall probe ("Tell me what happened"), rather than in response to direct suggestions, forced-choice questions, or limited-option questions.

It is true that children tend to be more accurate when asked open-ended questions that allow them to freely recall, compared to when they are asked more directed questions; this fact has been recognized since the earliest studies on children's suggestibility (Ceci & Bruck, 1995). However, it is not the case that children's spontaneous statements are always accurate. When children have been questioned suggestively, the suggestions can taint not only the statements made during that interaction but also the reports in later suggestion-free interactions, as seen in the Chester the Janitor study, described earlier. For example, the children in Poole and Lindsey's (2001) "Mr. Science" study, discussed earlier, were simply asked to describe everything they remembered from interacting with Mr. Science after they had been exposed to a misleading narrative by their parents about the event. A full 21% of the children's spontaneous statements had not been experienced, but had been suggested by their parents at a prior time. Furthermore, some of these events included bodily touch, such as Mr. Science putting something yucky in the child's mouth. In a different study, Poole and White (1993) interviewed children about an event that had occurred 2 years previously. Following the first event the children had been exposed to repeated questions, some of which were misleading. Two years later 39% of 6-year-olds' (who were 4 years old at the time of the original event) statements in response to the open-ended request, "Tell me what happened," were incorrect. Older children did not fare that much better, with 23% of 8-year-olds' and 25% of 10-year-olds' statements being incorrect.

Furthermore, even if children have not been exposed to misleading suggestions, their open-ended recall is not guaranteed to be accurate. This is especially true if the child is interviewed about a confusing or ambiguous event. For example, Ornstein and colleagues (1998) engaged children in a mock medical exam in which some common features (e.g., listening to the child's heart) were omitted and atypical features (e.g., wiping the child's belly button with alcohol) were added. When children were interviewed about the event 12 weeks later, 42% of 4-year-olds and 74% of 6-year-olds spontaneously reported at least one of the nonexperienced common features. In another study, Goodman and her colleagues (1994) interviewed children about a painful genital catheterization procedure. Among the children who were 3–4 years old, 23% of their free-recall statements were incorrect. Finally, Greenhoot (2000) examined 5- and 6-year-olds' recall of ambiguous stories, where the protagonist may have been considered acting either prosocially or antisocially. Error rates on open-ended recall ranged between 20 and 30%.

Taken together, these studies (and others reviewed by Kulkofsky et al., 2008) show that although children's free recall and spontaneous statements are generally more accurate than their responses to directed questions, free recall is by no means error free. Errors can result from suggestive techniques, but they can also reflect other factors that distort memory (e.g., forgetting, use of schemas to reconstruct theme-consistent but inaccurate details, misunderstanding). These distortions can turn into quite detailed false reports if the child is interviewed by a biased interviewer who conducts the interview to confirm the child's initial inaccurate statement. In contrast, an unbiased interviewer may be able to question the child in such a way as to test various hypotheses about the origin of the child's statements.

False Reports Produced by Suggestive Interviews Are Distinguishable from Accurate Reports

Expert witnesses often state that it is easy to detect false reports that are the product of false suggestions because such children (1) merely "parrot" the words of their interrogators, (2) sound rehearsed, and (3) provide unemotional narratives that have no sensory information (e.g., "It hurt"). These beliefs at times reflect confusion about reports that emerge as a result of suggestions, coaching, or lying. Coaching differs from suggestibility in that it implies a deliberate attempt on the part of the interviewer to put words into the child's mouth and to have the child actively rehearse these statements. In contrast, there is no deception involved when children provide

false reports in suggestive interviews. Children's inaccurate responses to the suggestions initially may reflect social pressure; they provide a response to please the interviewer. However, with time children come to believe that the falsely suggested event had actually happened. In other words, they develop a false belief about a statement they initially realized was false. In addition, deliberate attempts to coach the child are rare. In most cases, the interviewer is not consciously trying to put words into the child's mouth, and there are rarely deliberate attempts to rehearse the child's false testimony.

Thus, one cannot use the same criteria for detecting lies as for detecting false beliefs that result from suggestions. Yi and Ceci (2007) found that adult judges assess these coached children's reports as less credible than those of children who produced false reports as a result of suggestion or of children whose reports were accurate. In their study, nine preschool children reported either accurate or inaccurate information as a result of being coached to lie or being exposed to misinformation. College students rated the credibility of each child's statements. Children who were inaccurate due to having been exposed to misinformation appeared every bit as credible as children who were accurate, and both appeared more credible than children who had been coached to lie. We should note that other research, which has not included a suggestibility condition, indicates that adult judges cannot easily discriminate lies from truths in children who have been coached (Leach, Talwar, Lee, Bala, & Lindsay, 2004; Talwar, Lee, Bala, & Lindsay, 2006).

There are other consistent findings. Children who make false reports after being suggestively interviewed appear highly credible to trained professionals in the fields of child development, mental health, and forensics (e.g., Ceci, Crotteau-Huffman, Smith, & Loftus, 1994; Ceci, Loftus, Leicht-man, & Bruck, 1994; Leichtman & Ceci, 1995). As was the case with Yi and Ceci's college raters, these professionals cannot reliably discriminate children whose reports are accurate from those whose reports are inaccurate as a result of suggestive interviewing techniques. The children who provided the false reports spoke sincerely and provided accounts laden with emotion and perceptual details. Part of the difficulty in discerning true from false reports in children is that when children are suggestively interviewed, their subsequent narratives may include false reports that go beyond what was suggested to them but that are consistent with the suggestions (e.g., Bruck et al., 1995, 2002). For example, Leichtman and Ceci (1995) found that when children were given false suggestions that a person was "clumsy," they later generalized this information to make false claims that this individual spilled, ripped, soiled, or broke things during his visit to their classroom, even though these behaviors had not been part of the interviewer's suggestions.

Another set of studies has searched for objective characteristics that might differentiate suggested true from suggested false narratives. The frequency of linguistic markers, such as the amount of elaboration or temporal connectivity, do not consistently differentiate true from false narratives that emerge as a result of repeated suggestive interviews (Bruck et al., 2002; Powell, Jones, & Campbell, 2003; Scullin, Kanaya, & Ceci, 2002). In the Bruck and colleagues (2002) study described earlier, analyses showed that narratives of false events actually contained more embellishments (including descriptions and emotional terms) and details than their narratives of true events. The false narratives also had more spontaneous statements than the true narratives and included more temporal connectors. The robustness of false narratives may emerge because when children are suggestively interviewed they often learn that the truth value of their statements is not important; rather, the amount of detail about events that are of interest to the interviewer is what counts. This sometimes results in children providing very elaborate and sometimes bizarre reports about nonexperienced events.

These findings show that reliability and credibility are orthogonal dimensions. Children may appear highly credible (or their interviews may have the characteristics of credible narratives) and yet their reports may be unreliable. Children can also be very reliable and yet not appear credible in the eyes of judges and jurors. Accordingly, one cannot use perceived credibility to judge reliability. However, in the forensic arena, it is crucial for those who make judgments about credibility (i.e., the court) to make such judgments based on the reliability of the report. In other words, the judgment must take into account the degree to which suggestive interviewing specifically, and interviewer bias more generally, were used to elicit the child's statements. The role of the expert is to educate the court about the risks of various techniques that can taint children's testimony. With this knowledge, it is up to the court to decide if the facts of the case reveal the presence of such tainting mechanisms and the degree to which they are important for assessing the child's credibility.

GAPS IN RELEVANT SCIENTIFIC KNOWLEDGE

Since the late 1980s, there has been tremendous interest and progress in the field of children's suggestibility. Nevertheless, there are still many large gaps in our knowledge; here, we provide our list of the most crucial ones. First, there is almost no work on suggestibility and the factors that distort the testimony of adolescents. Given the great changes in the emotional, social,

neurobiological, and cognitive development of adolescents, it does not seem appropriate to treat them as big children or even as little adults. Compared with younger children and college students, adolescents may display very different mechanisms and motives for false allegations. One possibility is that they may be more sensitive to peer-group influences that would sway the accuracy of their testimony.

Second, although there are a number of good protocols for interviewing children who are victims of, or participants in, crimes, the training of interviewers does not specifically focus on the concept of "interviewer bias." Rather, existing guidelines mainly incorporate strategies that may decrease the number of suggestive techniques in an interview. Consequently, one could conduct a suggestion-free interview but end up with statements that, when put together, do not make sense or present a coherent narrative. This incoherence occurs because interviewers focus so much on the elements of the protocol and on the elicitation of single details that they never step back and ask, "Does this make sense?" or "What is this child trying to tell me?" This line of problem solving involves constructing a set of alternative hypotheses "online"—indeed a difficult task, but one that must be deconstructed and taught to obtain the best information from children.

Third, although there is now a substantial body of research dealing with individual differences among children that make some prone to suggestions and others highly resistant (e.g., Bruck & Melnyk, 2004; Chae & Ceci, 2005; Roebers & Schneider, 2001), there are virtually no data on individual differences that might render one better able to conduct a developmentally appropriate interview while avoiding interviewer bias. Although it would not be surprising if there were large differences among prospective trainees in the ability to multitask (e.g., keep track of the child's utterances, bearing in mind the case details, posit new hypotheses online), the more interesting and next question to address is the degree to which training can minimize such differences.

CONCLUSIONS

Scientifically Supported Uses of Interviewing

Children must be questioned by unbiased interviewers whose major task is to collect "untainted" evidence. If the child makes abuse-consistent statements, the interviewer must test alternative hypotheses about why the child is making that statement. Children's statements need to be elicited through open-ended questions or nonsuggestive techniques.

Scientifically Unsupported and Controversial Uses of Interviewing

We reviewed a number of suggestive techniques at the beginning of this chapter. When these techniques are used—especially by biased interviewers—the child may make abuse-consistent statements. The problem is that one cannot establish with any certainty whether the child's statement is accurate or false as a consequence of the suggestive techniques. In other words, the child's statements may have been tainted. Furthermore, there is no valid "Pinocchio test" that allows one to decide with any certainty whether a child's statements are accurate or false. Although suggestive interviewing techniques may still be controversial in some quarters, they are not supported by scientific evidence. Nevertheless, as we noted earlier, interviewer bias can distort the results of childhood interviews even in the absence of suggestive techniques, such as leading questions.

COMMUNICATING RELEVANT SCIENCE IN THE COURTROOM AND IN WRITTEN REPORTS

In our experience in the courtroom, the content of the material described in this chapter has been admissible under the *Daubert* standards. The studies have been published, the findings have been replicated, and there is general agreement in the field about most of these principles. Furthermore, based on surveys of potential jurors, this information is unfamiliar to laypersons.

It is always important to make sure that the court understands that this type of testimony speaks to the reliability (i.e., the trustworthiness) of the evidence, not to the honesty or credibility of a witness. Statements or reports can be unreliable due to normal processes of forgetting, distortion, and reconstruction. Statements can also be rendered unreliable if they are elicited in certain suggestive contexts. Thus, this type of testimony focuses on factors that enhance or degrade the quality of children's and adults' reports. Although this testimony does not assess the credibility (believability) of the child's allegations of abuse, information concerning the reliability of the child's report is crucial (for the court and nonexperts) to draw conclusions about the credibility of the children's allegations.

We have written a number of documents for the court that have presented the relevant scientific evidence and used it to provide possible explanations for the statements of the children in a particular case (e.g., see *Fuster Escalona v. Singletary; www.oranous.com/innocence/FrankFuster/MaggieBruck.html*

New Jersey v. Michaels, www.falseallegations.com/amicus.htm). There are some common themes in our written and oral testimonies. First, we only agree to educate the triers of fact in cases in which it is very clear that there was suggestive interviewing. We do not accept cases where there is little evidence or where we are asked to present a "theory" of why a child *might* make false statements. Consequently, on many occasions we have declined to participate because we found no evidence of taint in the case materials. In fact, in a few cases we worked with, and testified for, the prosecution regarding the clear absence of taint.

Second, our analysis ultimately examines the degree to which interviewer bias may have tainted the evidence in the case. This aspect involves more than examining the suggestive techniques during interviews, but extends to the whole conduct of the investigation; so this type of testimony is not always child-specific. For example, interviewer bias can be observed in a specific interview or by how an interviewer assesses the content or the impact of specific interviews. At a higher level, interviewer (or investigator) bias is revealed by the types of evidence that are pursued or ignored at the investigational level. For example, in many cases important leads are not followed up (because the primary and only hypothesis is that the prime suspect was the perpetrator).

Finally, although experts in this field cannot provide expert opinions as to the guilt of the alleged perpetrator or to the accuracy of the child's statements, they can provide an expert opinion concerning the amount of suggestion and taint present in the case. They can then conclude that, based on the scientific literature, this amount of taint or suggestion renders the child's statements unreliable.

REFERENCES

Brainerd, C. J., & Reyna, V. F., & Ceci, S. J. (2008). Developmental reversals in false memory: A review of data and theory. *Psychological Bulletin, 134,* 343–382.

Bright-Paul, A., Jarrold, C., & Wright, D. (2005). Age-appropriate cues facilitate source-monitoring and reduce suggestibility in 3- to 7-year-olds. *Cognitive Development, 20,* 1–18.

Bruck, M., Ceci, S. J., & Francoeur, E. (2000). A comparison of three and four year old children's use of anatomically detailed dolls to report genital touching in examination. *Journal of Experimental Psychology: Applied, 6,* 74–83.

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., & Hembrooke, H. (2002). Nature of true and false narratives. *Developmental Review, 22,* 520–554.

Bruck, M., Ceci, S. J., Melnyk, L., & Finkelberg, D. (1999, April). *Does interview bias create tainted reports?* Paper presented at the biannual meeting of the Society for Research in Child Development, Albuquerque, NM.

Bruck, M., Ceci, S. J., & Principe, G. (2006). The child and the law. In W. Damon & R. Lerner (Gen. Eds.), and K. A. Renninger & I. E. Sigel (Vol. Eds.), *Handbook of child psychology: Vol. 5. Child psychology in practice* (6th ed., pp. 776–816). New York: Wiley.

Bruck, M., London, K., Landa, R., & Goodman, J. (2007). Autobiographical memory and suggestibility. *Development and Psychopathology, 19,* 73–95.

Bruck, M., & Melnyk, L. (2004). Individual differences in children's suggestibility: A review and synthesis. *Applied Cognitive Psychology, 18,* 947–996.

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., Crotteau-Huffman, M., Smith, E., & Loftus, E. W. (1994). Repeatedly thinking about nonevents. *Consciousness and Cognition, 3,* 388–407.

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

Ceci, S. J., Papierno, P. B., & Kulkofsky, S. (2007). Representational constraints on children's suggestibility. *Psychological Science, 18*(6), 503–509.

Ceci, S. J., Ross, D., & Toglia, M. (1987). Suggestibility of children's memory: Psycholegal implications. *Journal of Experimental Psychology: General, 116,* 38–49.

Chae, Y., & Ceci, S. J. (2005). Individual differences in children's recall and suggestibility: The effect of intelligence, temperament, and self-perceptions. *Applied Cognitive Psychology, 19,* 383–407.

Connolly, D., & Price, H. (2006). Children's suggestibility for an instance of a repeated event versus a unique event: The effect of degree of association between variable details. *Journal of Experimental Child Psychology, 93*(3), 207–223.

Finnila, K., Mahlberga, N., Santtilaa, P., Sandnabbaa, K., & Niemib, P. (2003). Validity of a test of children's suggestibility for predicting responses to two interview situations differing in their degree of suggestiveness. *Journal of Experimental Child Psychology, 85,* 32–49.

Garven, S., Wood, J. M., & Malpass, R. S. (2000). Allegations of wrongdoing: The effects of reinforcement on children's mundane and fantastic claims. *Journal of Applied Psychology, 1,* 38–49.

Goodman, G. S., Quas, J. A., Batterman-Faunce, J. M., Riddlesberger, M., &

Kuhn, J. (1994). Predictors of accurate and inaccurate memories of traumatic events experienced in childhood. *Consciousness and Cognition, 3*, 269–294.

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

Greenhoot, A. F. (2000). Remembering and understanding: The effects of changes in underlying knowledge on children's recollections. *Child Development, 71*, 1309–1328.

Holliday, R. (2003). The effect of a prior cognitive interview on children's acceptance of misinformation. *Applied Cognitive Psychology, 17*, 443–457.

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.

Kulkofsky, S. C., Wang, Q., & Ceci, S. J. (2008). Do better stories make better memories?: Narrative quality and memory accuracy in preschool children. *Applied Cognitive Psychology, 21*, 21–38.

Lamb, M. E., Sternberg, K. J., & Esplin, P. W. (2000). Effects of age and delay on the amount of information provided by alleged sex abuse victims in investigative interviews. *Child Development, 71*, 1586–1596.

Lamb, M. E., Sternberg, K. J., Orbach, Y., Esplin, P. W., Stewart, H., & Mitchell, S. (2003). Age differences in young children's responses to open-ended invitations in the course of forensic interviews. *Journal of Consulting and Clinical Psychology, 71*(5), 926–934.

Leach, A., Talwar, V., Lee, K., Bala, N., & Lindsay, R. (2004). "Intuitive" lie detection of children's deception by law enforcement officials and university students. *Law and Human Behavior, 28*(6), 661–685.

Leichtman, M. D., & Ceci, S. J. (1995). The effects of stereotypes and suggestions on preschoolers' reports. *Developmental Psychology, 31*(4), 568–578.

London, K., Bruck, M., Ceci, S. J., & Shuman, D. (2005). Disclosure of child sexual abuse: What does the research tell us about the ways that children tell? *Psychology, Public Policy and Law, 11*, 194–226.

London, K., Bruck, M., Wright, D., & Ceci, S. J. (2008). Review of the contemporary literature on how children report sexual abuse to others: Findings, methodological issues, and implications for forensic interviewers. *Memory, 1*, 29–47.

Lyon, T. D. (2007). False denials: Overcoming methodological biases in abuse disclosure research. In M. E. Pipe, M. E. Lamb, Y. Orbach, & A. C. Cederborg (Eds.), *Disclosing abuse: Delays, denials, retractions, and incomplete accounts* (pp. 41–62). Mahwah, NJ: Erlbaum.

Malloy, L. C., Lyon, T. D., & Quas, J. A. (2007). Filial dependency and recantation

of child sexual abuse allegations. *Journal of the American Academy of Child and Adolescent Psychiatry, 46*(2), 162–170.

Marche, T. A. (1999). Memory strength affects reporting misinformation. *Journal of Experimental Child Psychology, 73,* 45–71.

Melnyk, L., & Bruck, M. (2004). Timing moderates the effect of repeated suggestive interviewing on children's suggestibility. *Applied Cognitive Psychology, 18,* 613–631.

Ornstein, P. A., Merritt, K. A., Baker-Ward, K., Furtado, E., Gordon, B. N., & Principe, G. (1998). Children's knowledge, expectation, and long-term retention. *Applied Cognitive Psychology, 12,* 387–405.

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

Poole, D. A., & Lindsay, D. S. (2001). Children's eyewitness reports after exposure to misinformation from parents. *Journal of Experimental Psychology: Applied, 7,* 27–50.

Poole, D. A., & White, L. T. (1993). Two years later: Effects of question repetition and retention interval on the eyewitness testimony of children and adults. *Developmental Psychology, 29,* 844–853.

Powell, M. B., Jones, C. H., & Campbell, C. (2003). A comparison of preschooler's recall of experienced versus non-experienced events across multiple interviewers. *Applied Cognitive Psychology, 17,* 935–952.

Powell, M. B., & Thomson, D. M. (1997). The effect of an intervening interview on children's ability to remember an occurrence of a repeated event. *Legal and Criminological Psychology, 2,* 247–262.

Reyna, V. F., & Brainerd, C. J. (1998). Fuzzy-trace theory and false memory: New frontiers. *Journal of Experimental Child Psychology, 71,* 194–209.

Roebers, C. M., & Schneider, W. (2001). Individual differences in children's eyewitness recall: The influence of intelligence and shyness. *Applied Developmental Science, 5,* 9–20.

Schreiber, N., & Parker, J. (2004). Inviting witnesses to speculate: Effects of age and interaction on children's recall. *Journal of Experimental Child Psychology, 89,* 31–52.

Scullin, M., Kanaya, T., & Ceci, S. J. (2002). Measurement of individual differences in children's suggestibility across situations. *Journal of Experimental Psychology: Applied, 8,* 233–246.

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

Sternberg, K. J., Lamb, M. E., Hershkowitz, I., Sternberg, K. J., Lamb, M. E., Hershkowitz, I., et al. (1996). The relation between investigative utterance types and the informativeness of the child witness. *Journal of Applied Developmental Psychology, 17,* 439–451.

Strange, D., Sutherland, R., & Garry, M. (2006). Event plausibility does not determine children's false memories. *Memory, 14*(8), 937–951.

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

Talwar, V., Lee, K., Bala, N., & Lindsay, R. C. L. (2006) Adults' judgments of child witness credibility and veracity. *Law and Human Behavior, 30,* 561–570.

Thompson, W. C., Clarke-Stewart, K. A., & Lepore, S. J. (1997). What did the janitor do?: Suggestive interviewing and the accuracy of children's accounts. *Law and Human Behavior, 21,* 405–426.

Yi, S., & Ceci, S. J. (2007, May). *Misinformation and deception detection among preschool children.* Paper presented at the annual meeting of the Association for Psychological Science, Washington, DC.