DOCUMENT RESUME

ED 390 517 PS 022 038

AUTHOR	Ceci, Stephen J.; Bruck, Maggie
TITLE	Child Witnesses: Translating Research into Policy.
INSTITUTION	Society for Research in Child Development.
SPONS AGENCY	National Inst. of Child Health and Human Development (NIH), Bethesda, Md.
PUB DATE	93
NOTE	33p.
AVAILABLE FROM	SRCD Executive Office, University of Michigan, 300 North Ingalls, 10th Floor, Ann Arbor, MI 48109-0406 ($3.50 for single issue; subscriptions available to nonmembers of SRCD at $12.50).
PUB TYPE	Collected Works - Serials (022)
JOURNAL CIT	Social Policy Report; v7 n3 Fall 1993
EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	Age Differences; *Child Abuse; Court Litigation; Incidence; Literature Reviews; Public Policy; *Reliability; *Sexual Abuse; *Social Science Research; *Young Children
IDENTIFIERS	Expert Witness; *Suggestion; *Testimony; Witnesses

ABSTRACT

This report provides an overview of the research on the testimony of young children in cases of sexual abuse, focusing on preschoolers' presumed suggestibility and the role of researchers and mental health professionals as expert witnesses in such cases. It does so in light of the McMartin preschool case, in which seven defendants were acquitted, after 7 years of legal proceedings, of having sexually abused a large number of preschoolers. The report examines the prevalence of child sexual abuse and the increasing acceptance by the courts of uncorroborated testimony by young children. Research is reviewed on the degree to which very young children are prone to suggestion, concluding that the most recent studies are more often ambivalent about the reliability of children's reports than earlier studies. However, recent studies suggest that there are reliable age differences in suggestibility, with preschoolers' reports more influenced by erroneous suggestions than older children's reports. Policy implications of suggestibility research are also discussed, including what expert witnesses should tell the court, the qualifications of expert witnesses, the relationship of research to clinical practice, and the role of professional organizations. (Contains 98 references.) (MDM)

SOCIAL POLICY REPORT

Society For Research In Child Development

Volume VII, Number 3

Fall 1993

CHILD WITNESSES: TRANSLATING RESEARCH INTO POLICY

Stephen J. Ceci
Maggie Bruck

On August 12, 1983, Judy Johnson, the mother of a toddler at a prestigious nursery school in Manhattan Beach, California, told police that her 2-year-old son had been molested by Raymond Buckey, a teacher and the grandson of the school's founder, Virginia McMartin. Buckey was arrested but subsequently released for lack of evidence. On March 22, 1984, he was indicted by a grand jury and rearrested along with six female teachers, including his mother, Peggy McMartin Buckey. He was held without bail until 1989.

In August, 1984, the first of many preliminary hearings had begun for the seven McMartin Preschool defendants. Fourteen former students at the nursery school took the witness stand at this hearing and described a series of bizarre events involving sexual abuse, satanic rituals, and animal mutilation that allegedly occurred at their preschool. Based on these children's testimony, the defendants were accused of 115 counts of abuse, later expanded to 321 counts, including rape, sodomy, fondling, oral copulation, and drugging, and photographing of at least 100 children in the nude.

In January 1986, after 17 months of preliminary hearings in which each child witness was cross-examined by each of seven different defense attorneys, charges against five of the six women were dropped because of insufficient evidence. Only 26-year-old Raymond Buckey and his 58-year-old mother, Peggy McMartin Buckey, remained as defendants. After spending 2 years in jail, Peggy McMartin Buckey's bail was set at $500,000 in 1986, while her son remained remanded without bail in a proceeding that would last an additional 4 years.

In the ensuing years of legal proceedings, the major issue before the court was whether or not to believe the children. On the one hand, it was argued that the children's reports were authentic and that their bizarre and chilling accounts of events, which were well beyond the realm of most preschoolers' knowledge and experience, only served to substantiate the fact that the children had

actually participated in them. On the other hand, it was argued that the children's reports were the product of repeated suggestive interviews by parents, law enforcement officials, and therapists, and that the children were only reporting events suggested to them during these interviews.

In January 1990, following a 33-month trial and 9 weeks of deliberation, the jury in the McMartin Preschool sexual abuse case returned "not guilty" verdicts on 52 of the 65 counts. The jury deadlocked on 12 molestation charges against Raymond Buckey and on 1 count of conspiracy against his mother. Judge William Pounders dismissed the conspiracy charge.

In response to the acquittals, the children's parents railed at the way the case had been handled and at the jury's verdicts. The children themselves appeared on nationally televised talk shows, weeping over the jury's seeming refusal to believe their claims of ritualistic abuse. Newspapers and magazines across the nation ran headlines such as: "Doubt the children and jail the parents!"

During a postverdict press conference, many jurors claimed that they believed that some of the children had been abused, but were unable to reach a guilty verdict because of the suggestive way the children had been interviewed. These jurors claimed that the social workers prevented the children from speaking in their own language and thus diminished their credibility.[1]

By January 1990, the prosecution announced Buckey would be retried on the remaining 12 charges on which the jurors had been deadlocked. Although this trial was shorter, the jury again deadlocked. When the judge declared a mistrial, the prosecution did not to retry the case.

The McMartin case made legal history. In sheer magnitude, it was without parallel, lasting 7 years from the time of its inception to the final verdicts, producing hundreds of thousands of pages of transcript, and costing the State of California over $16 million. From the very start, the McMartin case captured the attention of the national media, with regular accounts of the children's allegations appearing on television (e.g., ABC News *Nightline*, April 20, 1984), in newspapers (e.g., Charlier & Downing, 1988; Shaw, 1990) and in magazine articles (e.g., Fischer, 1988).

The McMartin case is not a singular happening. There have been many similar cases in North America and Europe, some of which have received extensive media attention.[2] For the most part, these cases share the following elements: First, the witnesses were preschoolers at the time of the alleged abuse. Second, the disclosures were not made immediately following the alleged event, but after a long delay. Third, the disclosures often were preceded by intensive interviewing of the children by various professionals (e.g., child protective service workers, law enforcement) and nonprofessionals (e.g., parents, grandparents). Fourth, the children were the only witnesses to these alleged events, and corroborative physical evidence was lacking. Fifth, none of the defendants ever made a confession; all maintained their innocence, even after some co-defendants were convicted. Finally, the major issue before the jury in all of these cases was whether to believe the children.

Researchers in child development have served as witnesses or consultants for the defense and/or the prosecution in all of these trials. More importantly, these cases have changed the course of our research and thinking on children's memory development. They have encouraged researchers to tackle new issues, to develop innovative experimental paradigms, and to challenge and elaborate previous research on the reliability of young children's statements.

In this *Social Policy Report* we provide a glimpse into the social science research that has

accumulated on the aspect of children's testimony that figured so prominently in the McMartin case and hundreds of others like it,[3] namely, preschoolers' presumed suggestibility. We then present some tentative thoughts about the policy implications of this research, by addressing the questions: What can (and should) researchers and mental health professionals tell courts when they are called upon to serve as expert witnesses and consultants? What is the proper role for professional organizations to play in overseeing expert testimony? First, however, we provide some background information about the problem of child sexual abuse in the U.S. and about the history of children providing courtroom testimony.

Prevalence of Abuse and Court Involvement

Crime statistics reflecting the sexual abuse of children are of great social concern. These statistics come from two major sources: one is based on annual rates of "substantiated" or "indicated" reports of child abuse,[4] and the other is based on adults' reports of abusive events during their own childhoods. The first source thus provides estimates of *incidence*, whereas the second provides estimates of *prevalence*.

According to the most recent *incidence* figures (based on data from 45 states), there were 2.7 million reports of suspected child maltreatment in 1991 (National Center for Child Abuse and Neglect, 1993); 129,697 of these were substantiated or indicated cases that were sexual in nature, indicating an incidence of childhood sexual abuse of less than 1% of children for that year. The accuracy of these rates has been challenged, (e.g., Besharov, 1991; Robin, 1991). On the one hand, these figures may overestimate the extent of child sexual abuse because they include indicated and substantiated cases which are not validated. Others, however, argue that these rates may underestimate the incidence of child sexual abuse, because many cases of actual child sexual abuse end up being classified "unsubstantiated," or significant numbers of cases of abuse are never reported to authorities. It seems plausible that while some substantiated cases are actually false, many more unsubstantiated and unreported cases are real. Thus, the national data likely underestimate the true incidence, although no one can say by how much.

Random samples of adults, asked about their childhood history, yield highly variable estimates of childhood sexual abuse, ranging, for females, from 6.8% (Siegel, Sorenson, Golding, Burnam, & Stein, 1987) to 62% (Wyatt, 1985); and for males, from 3% to 31% (see Peters, Wyatt, & Finkelhor, 1986). A number of methodological factors may account for these discrepant figures, but review of these is beyond the scope of this article. Our point is that even if the lowest prevalence rates are the most accurate, and even if the incidence of child sexual abuse in 1991 was 1%, this still represents a serious societal problem.

As a result of society's reaction to these dramatic figures, and particularly in reaction to the ineffective prosecution of child abuse cases, the legal system has been forced to change some of its rules concerning the admissibility and treatment of child witnesses. For example, until recently courts of law in all English-speaking countries were reluctant to accept the uncorroborated statements of child witnesses (Chadbourn, 1978). This reluctance was reflected in competency hearings, corroboration requirements, and cautionary instructions that some judges gave to juries concerning the risk of convictions based solely on the testimony of child witnesses (Andrews, 1964; Cohen, 1975). During the 1980s, however, all but a few states dropped their corroboration requirement for children in sexual abuse cases, a crime that by its nature lacks corroboration. Seventeen

states now allow children to testify regardless of the nature of the crime, permitting the jury to determine how much weight to give to the child's testimony. With the continued adoption by states of Federal Rules of Evidence, the number of child witnesses is likely to expand.

As more and more children have been admitted as witnesses in the courtroom, legal and courtroom procedures have been modified. For example, some courts have instituted shield laws which permit a child witness to testify either behind a one-way screen or over closed-circuit television, to occlude the child's view of the defendant but not the defendant's view of the child. Hearsay exceptions are also allowed, whereby therapists, pediatricians, and others describe what children have said to them. These measures serve to assist child witnesses who otherwise might be "psychologically unavailable" to testify in open court (McGough, in press; Montoya, 1992, 1993).

In light of claims that many of these modifications challenge the constitutional rights of defendants (*Maryland* v. *Craig*, 1990; *Coy* v. *Iowa*, 1988), it is important to determine whether such procedures do, in fact, facilitate the accuracy of children's testimony (Montoya, 1993). This is particularly important in light of the fact that recent court decisions regarding the treatment of child witnesses have not been predicated exclusively on humane issues (i.e., reducing the stress placed on child witnesses), but also on the presumption that courtroom modifications will increase testimonial accuracy (Harvard Law Review Notes, 1985; Montoya, 1992). No scientific data addressing these issues are yet available, although some data on the costs and benefits of courtroom innovations on children's courtroom behavior have been gathered (Batterman-Faunce & Goodman, 1993; Flin, 1993).

There are, however, no reliable national data on the impact of these changes on the number of children who actually end up participating in family court or criminal justice proceedings. Gray's recent analysis (1993) of eight jurisdictions suggests that 3% to 10% of all cases of sexual abuse that are eventually filed with police result in a trial. Based on the National Center on Child Abuse and Neglect (1993) statistics cited above, this would suggest that, in the 45 states reporting, up to 13,000 children testified in sexual abuse trials that year. For various reasons, however, the actual incidence of court involvement may be considerably higher. The majority of cases that end in pleas still require the child to be deposed even if the child does not testify in a court trial. Second, these figures do not include data from five states, one of which is New York where as many as 3,150 children had formal court involvement in sexual abuse cases in 1990 (Doris, 1993). Finally, if nonsexual types of abuse and nonabuse cases involving children serving as witnesses are included (e.g., cases of domestic violence, custody disputes, accidents, playground injuries), then the estimate of children's participation in the legal system rises considerably, possibly to over 100,000 cases annually. We must emphasize, however, that this is, at best, an educated guess.

One final point about children's court involvement is in order. It appears that preschoolers are disproportionately more likely to be abused and more likely to have their cases come to trial. In an analysis of a sample of nearly 800 alleged victims of child sexual abuse in New York, preschoolers (ages 6 and younger) accounted for nearly 40% of the official sexual abuse cases, and 28% were aged 5 and younger (Doris, 1993). In Gray's recent analysis (1993), children below the age of 8 accounted for 45% of sexual abuse cases, and 18% were 5 years old or younger; 31% of the cases involving 5- to 6-year-old children went to trial, and 24% of cases involving 3- to 4-year-olds

went to trial, whereas only 10% of the cases involving 13- to 14-year-olds went to trial.

Despite modifications in the judicial system resulting in the greater court involvement of children, both jurists and social scientists continue to raise fundamental questions about whether these changes actually facilitate the accuracy of children's testimony (Montoya, 1992; 1993) and, more broadly, whether the testimony that children do give is accurate. Next, we discuss the research that has been carried out on one important aspect of the accuracy of children's reports: the degree to which very young children are disproportionately prone to suggestion.

Research on Children's Suggestibility: Past and Present Trends

The scientific research on the suggestibility of children's recollections is both contradictory and confusing. A review of 20th-century studies of children's suggestibility can be found in Ceci & Bruck, 1993. Our purpose in this report is not to recap that analysis but to highlight some of the salient conclusions of these studies and to focus on the different experimental approaches used in investigating children's suggestibility.

Early Studies of Suggestibility

Early studies of children's suggestibility, with few exceptions, led to a jaundiced portrayal of children's proneness to suggestibility. Beginning with the early experiments of Binet and his European colleagues (Binet, 1900; Lipmann, 1911; Stern, 1910; Varendonck, 1911), and concluding with empirical studies in the 1920s and 1930s (Messerschmidt, 1933; Otis, 1924; Sherman, 1925), early researchers viewed children as extremely susceptible to leading questions and unable to resist an interviewer's suggestions. M. R. Brown (1926), a legal scholar, wrote:

> Create, if you will, an idea of what the child is to hear or see, and the child is very likely to see or hear what you desire. (p. 133)

Although the conclusions of these early researchers were confirmed by studies conducted right up until the 1980s, modern researchers have been ambivalent about generalizing these results to the forensic arena for several reasons. First, despite the fact that there is great concern currently about the reliability of preschoolers' reports, not one study in the first 80 years of this century included preschoolers. More recent research has begun to fill this void; since 1980, over 20 studies relevant to the issues of children's suggestibility have included a preschool sample.

A second and more important concern was that most of the previous studies involved children's recall of events that were forensically irrelevant. In most of this earlier literature, researchers examined the influences of a single misleading suggestion or a leading question on children's reports of neutral, nonscripted, and often uninteresting events that occurred in a laboratory setting. Although these results may be of importance for theoretical conceptualizations of the mechanisms that underlie suggestibility effects and memory processes, they have limited practical and legal relevance to the reliability of the child witness. In many court cases, the allegations involve the child as participant and not as bystander; they involve the child's recall of salient, rather than peripheral, events; they often involve repeated interviews which are highly suggestive; and they frequently involve emotionally charged and highly stressful events, such as sexual molestation. The earlier experiments of this century provide no clues as to the testimonial accuracy of children in such circumstances.

But how does a researcher conduct an ethically acceptable experiment that mirrors the many

conditions that are characteristic of the child victim–witness? It would be unacceptable, for example, to determine whether an interviewer can successfully suggest to children with substantiated histories of abuse that the abuse had never taken place. Similarly, it would be unacceptable to determine if nonabused children will make allegations of sexual abuse after a highly suggestive interview. It is ethically impermissible to alter such fundamental aspects of young children's autobiography.

Modern Studies of Children's Suggestibility

In the past several years, a number of researchers have attempted to deal with these issues by developing new paradigms which admittedly do not mirror all of the conditions that bring children to court, but which do contain some important elements of the child witness's experiences. This section describes three major lines of recent research, each of which illustrates a different paradigm: (1) increasing the salience of the experienced events about which children will be interviewed, (2) increasing the dynamics of the interview situation, and (3) adding anatomically correct dolls to the interviewing context.

Increasing the salience of events. As discussed above, earlier studies were criticized as not forensically relevant because they did not examine how children respond to questions about events that involved their own body, or about other salient events that occurred in personally experienced and stressful situations. In response, a number of researchers have designed studies in which children are asked misleading questions about being touched. In some studies, children are questioned about their previous interactions with an experimenter in a laboratory (e.g., Rudy & Goodman, 1991). In other studies, children are questioned about an inoculation (Goodman, Hirschman, Hepps, & Rudy, 1991) or a genital examination (Saywitz, Goodman, Nicholas, & Moan, 1991).

For example, Saywitz and her colleagues (Saywitz et al., 1991) examined 5- and 7-year-old girls' memories of a medical examination. Half of each age group had a scoliosis exam (for curvature of the spine), and the other children had a genital exam. When children were interviewed either 1 or 4 weeks later, they were asked suggestive and nonsuggestive questions that were abuse-related (e.g., "How many times did the doctor kiss you?") or nonabuse-related (e.g., "Didn't the doctor look at your feet first?"). Although the older children were initially more accurate than the younger children on most questions, some of these age differences disappeared after the 4-week delay. Most importantly, although there were age differences in response to the suggestive abuse questions, very few children of either age gave incorrect responses; the 7-year-old children never made a false report of abuse, and the 5-year-olds did so only 4 times, although they were given 215 opportunities.

Saywitz and her colleagues point out specific patterns of results in this study. They conclude the children's inaccurate reports involved mainly errors of omission rather than commission. The majority of children in the genital examination condition did not disclose genital contact unless specifically asked to do so. This latter opportunity was only provided with the direct (leading) question format ("Did the doctor touch you here?"). In the scoliosis condition, when children were asked these direct questions, 2.86% of the children falsely affirmed vaginal touch and 5.56% falsely affirmed anal touch.[5] In reviewing this study, Goodman and Clarke-Stewart (1991) conclude that:

> . . . obtaining accurate testimony about sexual abuse from young children is a complex task. Part of the complexity rests in the fact that

> there are dangers as well as benefits in the use of leading questions with children. The benefits appear in the finding . . . that leading questions were often necessary to elicit information from children about actual events they had experienced (genital touching). . . . The children. . . . were generally accurate in reporting specific and personal things that had happened to them. If these results can be generalized to investigations of abuse, they suggest that normal children are unlikely to make up details of sexual acts when nothing abusive happened. They suggest that children will not easily yield to an interviewer's suggestion that something sexual occurred when in fact it did not, especially if nonintimidating interviewers ask questions children can comprehend. (pp. 102-103)

Thus, according to this group of researchers, earlier studies of children's suggestibility may have overestimated the extent to which they are suggestible. For example:

> There is now no real question that the law and many developmentalists were wrong in their assumption that children are highly vulnerable to suggestion, at least in regard to salient details. Although some developmentalists may be challenged to find developmental differences in suggestibility in increasingly arcane circumstances, as a practical matter who really cares whether 3-year-old children are less suggestible about peripheral details in events that they witnessed than are 4-year-old children? Perhaps the question has some significance for developmental theory, but surely it has little or no meaning for policy and practice in child protection and law. (Melton, 1992, p. 154)

It is important, however, to point out that not all data on children's reports of medical procedures are consistent with these conclusions. Ornstein and his colleagues (Baker-Ward, Gordon, Ornstein, Larus, & Clubb, in press; Ornstein, Gordon, & Larus, 1992) found that when children were later questioned about their memories of the visit to the pediatrician, 3-year-olds were more prone than 6-year-olds to make false claims in response to suggestive questions about silly events involving body contact (e.g., "Did the nurse lick your knee?"). Oates and Shrimpton (1991) also found that preschoolers were more suggestible than older children about previously experienced events that involved body touching. In contrast to the Saywitz et al. findings that false reports in response to suggestive questions are relatively infrequent, the younger children in these latter studies provided a substantial number of false reports in response to suggestive questions. Until recently, however, only a few studies have included explicit questions about sexual touching. Recent research by us and our colleagues has yielded different results, which will be reported in greater detail later when we describe our study of a pediatric examination.

Increasing the dynamics of the interview. A second major innovative theme in the current research on children's suggestibility involves examining the effects of various interviewing techniques on children's reports. This focus has arisen in response to the concern that the interviewing procedures of earlier studies were less intense than

those that bring children to court—so much so as to result in a potential *underestimation* of children's suggestibility (Raskin & Esplin, 1991; Steller, 1991).

The interviewing procedures used in traditional laboratory studies and those used in the forensic arena differ in several ways. First, it is frequently the case that children who come to court are questioned weeks, months, or even years after the occurrence of an event, as opposed to minutes or days later. Suggestibility effects may be more salient after long delays, because the original memory trace has faded sufficiently to allow the suggestion to intrude more readily than might occur after shorter delays.

Second, child witnesses are rarely interviewed only one time, by one interviewer, or under nonstressful conditions. The modal child witness has been interviewed between 4 and 11 times prior to the first courtroom appearance; sometimes children are interviewed weekly for years about the same event—in therapy sessions, for instance. Leichtman and Ceci (1993) have suggested that the incessant use of leading questions and suggestions in these repeated interviews may result in a qualitatively different type of report distortion than that which arises from a single misleading question in a single postevent interview.

Third, an examination of the interviews of some child witnesses reveals that the label "suggestive interview" may describe more than the use of misleading questions. Rather, implicit and explicit suggestions can be woven into the fabric of the interview through the use of bribes, threats, repetitions of certain questions, and the induction of stereotypes and expectancies (Ceci & Bruck, 1993).

Finally, the questioning of child witnesses is typically conducted by parents, therapists, and legal officials, all of whom represent status and power in the eyes of the child; children may thus be more likely to comply with the suggestions of these interviewers than with those of the neutral interviewers employed in most research studies.

Although it is very difficult to create experimental conditions that simulate the confluence of the conditions present in child witness interviews (stressful episodes, with repeated and suggestive questioning over prolonged periods of time), researchers are beginning to examine how children's reports are influenced by the repetition of suggestions in multiple interviews prior to and following the occurrence of an event. In addition, researchers have focused on the interviewer and the potential effects that a particular interviewer's bias may have on the reports elicited from young children. We confine our discussion here to three studies recently carried out with our colleagues at Cornell and McGill universities, as they were designed specifically to address these issues (see Ceci & Bruck [1993], for discussion of additional studies). In focusing on our own studies, we necessarily present a particularized view, inspired by our own hypotheses, assumptions, and values, but the studies are designed to build on and challenge extant research.

In these studies, we patterned our experimental manipulations after materials collected over the past decade from court transcripts and from therapy sessions and law enforcement interviews involving children in cases similar to the McMartin case where there was a strong suspicion of abuse (see transcripts in Ceci, in press; Ceci & Bruck, 1993). These materials reveal that a child's first "disclosure" about abuse commonly occurs when an interviewer pursues a single hypothesis about the basis of the child's difficulties, which entails leading and suggestive interviews, often with fantasy inductions and "self-empowerment" techniques– the techniques themselves being poten-

tially suggestive and stereotype inducing. Such disclosures are then pursued in law enforcement, child protective service, or therapeutic interviews.

Study 1: The effect of interviewer bias on children's reports

Ideally a forensic interview should be guided by a hypothesis-testing framework. Just as scientists try to arrive at the truth by ruling out rival hypotheses or by falsifying a favored hypothesis (Ceci & Bronfenbrenner, 1991; Dawes, 1992; Popper, 1962), interviewers should, in similar manner, attempt to rule out rival hypotheses, rather than exclusively attempting to confirm their favored one. However, because of situational pressures (e.g., case workers must sometimes make immediate determinations of potential danger to a child), it is not feasible that interviewers generate and test every conceivable hypothesis or, conversely, that they be "blind" to obviously relevant information pertaining to a main hypothesis that abuse is indeed present. Failure to recognize relevant information provided by the child could result in crucial missed opportunities.[6] But, as the following study shows, failure to test a rival hypothesis can result in reporting errors.

In this study (Ceci, Leichtman, & White, in press), we examined how an interviewer's hypothesis can influence the accuracy of young children's reports. Preschoolers were exposed to a game-like event and then interviewed 1 month later. The interviewer was given some information about events that *might* have occurred; some of the information was accurate and some of it was inaccurate. The interviewer was told to interview each child and to use whatever strategies she felt necessary to elicit the most factually accurate report from the child. The information we provided influenced the interviewer's hypotheses about what had transpired in this game, which, in turn, appeared to exercise a powerful influence on the dynamics of the interview, with the interviewer eventually shaping some of the children's reports to be consistent with her hypothesis about what had happened. When the interviewer was accurately informed, she got children to recall correctly 93% of the events that had transpired. It is important to note that the children made no false accusations when the interviewer was correctly informed, that is, they only made "errors of omission." However, when the interviewer was misinformed, 34% of the 3- to 4-year-olds and 18% of the 5- to 6-year-olds corroborated one or more false events that the interviewer erroneously believed had transpired. Thus, in the misinformed condition, the children made "errors of commission." Finally, the children seemingly became more credible as the interview unfolded. Many children initially stated details inconsistently, or with reluctance or even denial, but as the interviewer persisted in asking about nonevents, some abandoned their hesitancy and denials.

Because the interviewers were trained professionals (one was an experienced social worker, the other a nursery school teacher), we feel that the types of interactions observed in this study may be similar to those that occur in interviews between young children and parents, teachers, and professionals who are not given explicit training in how to generate and test alternative hypotheses. Our review of the materials from some publicized cases, such as McMartin, reveals that professional interviewers often steadfastly stick with one line of inquiry even when children continue to deny that the questioned events ever occurred (for examples see Ceci, in press).

Study 2: The effects of stereotype induction and repeated suggestions on young children's reports

A stranger named Sam Stone paid a 2-minute visit to preschoolers (aged 3 to 6) in their day-care

center (Leichtman & Ceci, in press). Following Sam Stone's visit, the children were asked for details about the visit on four different occasions over a 10-week period. On each occasion, the interviewer refrained from using suggestive questions; she simply encouraged children to describe Sam Stone's visit in as much detail as possible. One month later, the children were interviewed a fifth time by a new interviewer, who first elicited a free narrative about the visit. Then, using probes, she asked about two "nonevents" which involved Sam Stone doing something to a teddy bear and a book. In reality, he never touched either item.

When asked in the fifth interview, "Did Sam Stone do anything to a book or a teddy bear?" most children accurately replied, "No." Only 10% of the youngest (3- to 4-year-old) children's answers contained claims that Sam Stone did anything to a book or teddy bear. When asked if they actually *saw* him do anything to the book or teddy bear, as opposed to "thinking they saw him do something," or "hearing he did something," now only 5% of their answers contained claims that anything occurred. Finally, when these 5% were gently challenged ("You didn't really see him do anything to the book/the teddy bear, did you?") only 2.5% still insisted on the reality of the fictional event. None of the older (5- to 6-year-old) children reported that they had seen Sam Stone do either of the fictional actions.

Another group of preschoolers was presented with a stereotype of Sam Stone before he ever visited their school. We did this to mimic the sort of stereotypes that some child witnesses have acquired about actual defendants. (In actual cases, for example, some children have been told repeatedly that the defendant did "bad things.") Each week, beginning a month prior to the visit, the children in our study were told a new Sam Stone story in which he was depicted as very clumsy. For example:

> You'll never guess who visited me last night. [pause] That's right. Sam Stone! And guess what he did this time? He asked to borrow my Barbie and when he was carrying her down the stairs, he tripped and fell and broke her arm. That Sam Stone is always getting into accidents and breaking things!

Following Sam Stone's visit, these children were interviewed four times over a 10-week period. These four interviews contained erroneous suggestions (e.g., "When Sam Stone ripped that book, was he being silly or was he angry?"). At the fifth interview, these children were asked for a free narrative about Sam's visit and were then asked probing questions about the two nonevents.

In this last interview, 72% of the youngest preschoolers claimed that Sam Stone did one or both misdeeds, a figure that dropped to 44% when they were asked if they actually saw him do these things. Importantly, 21% continued to insist that they saw him do these things, even when gently challenged. The older preschoolers, though more accurate, included 11% of children who insisted they saw him do the misdeeds.

Some researchers have opined that the presence of perceptual details in reports is one of the indicators of an actual memory, as opposed to a confabulated one (Raskin & Yuille, 1989; Schooler, Gerhard, & Loftus, 1986). In this study, however, the presence of perceptual details was no assurance that the report was accurate. In fact, children in the stereotype plus suggestion condition produced a surprising number of fabricated perceptual details to embellish their false accounts of nonevents (e.g., claiming that Sam Stone took the teddy bear into a bathroom and soaked it in hot water before smearing it with a crayon). The difference in the quality of reports obtained in this study compared to others in the suggestibility

literature may reflect the conditions under which the reports were obtained. As mentioned earlier, in most past studies, children's erroneous reports were in response to a single misleading question, posed after a brief delay following the event in question. In contrast, in the present study, children's false reports were a product of repeated erroneous suggestions over a relatively long period of time, coupled with a stereotype that was consistent with these suggestions.

It is one thing to demonstrate that children can be induced to make errors and include perceptual details in their reports, but it is another matter to show that their faulty reports are convincing to others. To examine the believability of the children's reports, we showed videotapes of their final interview to approximately 1,000 researchers and clinicians who work on children's testimonial issues. These researchers and clinicians were told that all the children observed Sam Stone's visit to their day-care centers. They were asked to decide which of the events reported by the children actually transpired and then to rate the overall credibility of each child.

The majority of the professionals were inaccurate. Analyses indicated that these experts—who conduct research on the credibility of children's reports, provide therapy to children suspected of having been abused, or carry out law enforcement interviews with children—generally failed to detect which of the children's claims were accurate, despite being confident in their judgments. Since so many of the children claimed that Sam Stone ripped the book and/or soiled the bear, it is understandable that many of the experts reasoned that these events must have transpired. But their overall credibility ratings of individual children were also highly inaccurate, with the very children who were least accurate being rated as most accurate. We believe that the highly credible yet inaccurate reports obtained from the children resulted from a combination of repeated interviews with persistent and intense suggestions that built on a set of prior expectations (i.e., a stereotype). In a similar way, it may become difficult to separate credibility from accuracy when children, after repeated interviews, give a formal videotaped interview or testify in court.

Study 3: Influencing children's reports of a pediatric visit

It could be argued that the Sam Stone Study is not relevant to evaluating the reliability of a child witness who reports personally experienced events involving his or her own body, especially when the experience involves some degree of distress. Furthermore, some might argue that the Sam Stone data are not germane to testimony about highly predictable and scripted events. In cases where the event involves a child's own body, is somewhat stressful, and is predictable, it is often thought that children may be less prone to suggestion.

To determine if children could be misled under such circumstances, we examined the influence of postevent suggestions on children's reports about a pediatric visit where they were examined (Ceci, Leichtman, & Bruck, in press). The study had two phases. In the first phase, 5-year-old children visited their pediatrician for an annual check-up. A male pediatrician examined the child. Then the child met a female research assistant who talked about a poster that was hanging on the wall in the examining room. Next, the pediatrician gave the child an oral polio vaccine and a DPT inoculation. Then the research assistant gave the child one of three types of feedback about how the child had acted when receiving the inoculation. One group was given pain-affirming feedback; they were told that it seemed as though the shot really hurt them, but shots hurt even big kids

(hurt condition). A second group was given pain-denying information; these children were told that they acted like the shot did not hurt much, and that they were really brave (no-hurt condition). Finally, a third group was merely told that the shot was over (neutral condition). After the feedback, the research assistant gave each child a treat and then read the child a story. One week later, a different assistant visited the children and asked each one to indicate through the use of various rating scales how much he or she had cried during the shot and how much the shot hurt.

The children's reports did not differ as a function of feedback condition. Thus, we found that children could not be influenced to make inaccurate reports concerning significant and stressful procedures involving their own bodies. These results are similar in spirit to those of Saywitz et al. (1991) who also provided children with suggestions about stressful, personally experienced events in a single interview and discovered that children can be quite resistant to erroneous suggestions.

In the second phase of our study, we reinterviewed the children three more times, approximately 1 year after the shot. During these interviews, children were provided with repeated suggestions about how they had acted when they received their inoculations. Thus, as in the first phase of the study, some children were told that they were brave when they got their shot, whereas other children were not given any feedback. (For ethical reasons, we provided only "no-hurt" and "neutral" feedback in this phase of the study. We felt that providing "hurt' feedback might induce false or unpleasant memories about visiting the doctor.) When the children were visited for a fourth time and asked to rate how much the shot had hurt and how much they had cried, there were large suggestibility effects. Those who had been repeatedly told that they had acted brave when they had received their inoculation a year earlier reported significantly less crying and less hurt than children who were given no feedback. Thus, these data indicate that children's reports of stressful events involving their own bodies can be distorted under certain circumstances.

In the second phase of this study, we also tried to mislead children about the people who performed various actions during the original inoculation visit. Some children were falsely reminded on three occasions that the pediatrician gave them treats, showed them the poster, and read them a story. Some children were falsely reminded on three occasions that the research assistant gave them the inoculation and the oral vaccine. Control children were merely reminded that "someone" did these things. Based on the conclusions of other researchers (e.g., Fivush, 1993; Melton, 1992), it was hypothesized that children should not be suggestible about such important events and that they should be particularly immune to suggestions that incorporate shifts of gender. The male pediatrician had never given them treats or read them a story, and the female research assistant had never performed any medical procedures.

Contrary to these predictions, the children were misled. In the fourth interview, when asked about their doctor's visit in the previous year, 67% of the children (versus 27% of the control children) who were given misleading information about the pediatrician reported that the pediatrician showed them the poster, gave them treats, or read them a story. For children who were falsely told that the research assistant had given them the shot and the vaccine, 50% (versus 16% of the control children) fell sway to at least one of these two suggestions. Interestingly, 38% of the children who were given misleading information that the research assistant gave them the oral vaccine and the inoculation also said that the research

assistant had performed other scripted events that not only had never occurred but also had never been suggested (e.g., reporting that the research assistant checked their ears and nose). None of the control children made such inaccurate reports. Thus, our suggestions influenced not only children's reports of personally experienced, salient events, but also their reports for nonsuggested scripted events that were related to the suggested events.

These data indicate that under certain circumstances children's reports concerning stressful events involving their own bodies can be influenced. The two factors that were most critical to this pattern of results were the intensity of the suggestions (i.e., repeating the suggestions over multiple interviews) and their timing (i.e., the long delay between the original event and interview about the event). These same two factors are characteristic of the conditions under which children made allegations of sexual abuse in many of the cases described at the beginning of this report.

The results of this study are consistent with the Sam Stone study even though the nature of the events about which children were misled were different. In the Sam Stone study, repeated suggestions and stereotypes led to convincing fabrications of nonoccurring events. In the pediatrician study, misleading information given in repeated interviews after a long delay following a target event influenced children's reports of personally-experienced, salient events.

The suggestibility of anatomically correct dolls. Anatomically correct dolls are frequently used by professionals, including child therapists, police, child protection workers, and attorneys, in interviewing children about suspected sexual abuse. According to recent surveys, 90% of field professionals use anatomical dolls at least occasionally in their investigative interviews with children suspected of having been sexually abused (Boat & Everson, 1988; Conte, Sorenson, Fogarty, & Rosa, 1991). Although no national figures are available, it appears that expert testimony is often based on observations of children's interactions with such dolls (Mason, 1991). We include a discussion here of anatomical dolls, because a number of commentators have raised questions about whether the dolls are suggestive (e.g., McGough, in press; Moss, 1988; Raskin & Yuille, 1989).

One rationale for the use of anatomical dolls is that they allow children to manipulate objects reminiscent of a sexual event, thereby cuing recall and overcoming language and memory problems. Another rationale is that their use is thought to overcome embarrassment and shyness. The dolls have also been used as projective tests. Some claim that if a child actively avoids these dolls, shows distress if they are undressed, or shows unusual preoccupation with their genitalia, this is consistent with the hypothesis that the child has been abused (see Mason, 1991).

The use of anatomically correct dolls has raised skepticism, however, among researchers and professionals alike. Two related arguments are frequently invoked against their use. The first is that the dolls are suggestive, that they encourage the child to engage in sexual play even if the child has not been sexually abused (e.g., Gardner, 1989; Terr, 1988). A child may insert a finger into a doll's genitalia, for instance, simply because of its novelty or "affordance," much the way a child may insert a finger into the hole in a doughnut. Another criticism is that it is impossible to make firm judgments about children's abuse status on the basis of their doll play because there are no normative data on nonabused children's doll play.

In several studies, researchers have compared the doll play of sexually abused and nonabused children. In addition, there have been a score of studies examining the doll play of

nonabused children. Reviews of this literature (Berry & Skinner, 1993; Ceci & Bruck, 1993; Wolfner, Faust, & Dawes, 1993) indicate that many of the studies are methodologically inadequate and do not allow for firm interpretations about the potential usefulness or risks of using dolls. Furthermore, some data indicate that some of the play patterns thought to be characteristic of abused children, such as playing with the dolls in a suggestive or explicit sexual manner, or showing reticence or avoidance when presented with the dolls, also occur in samples of nonabused children (see Bruck & Ceci [1993] for a review). Finally, other data indicate that the dolls, though not suggestive, do not improve reporting—particularly among younger children (e.g., Goodman & Aman, 1990).

We have recently completed a study of 3-year-old children's interactions with anatomically correct dolls that highlights each of these results (Bruck, Ceci, Francœur, & Renick, in press). The children in this study visited their pediatrician for their annual check-up. The pediatrician conducted genital examinations with half the children; the remaining children did not receive genital exams. Immediately after the examination, the child was interviewed by a research assistant. Pointing to the buttocks and then to the genital areas of an anatomically correct doll, the assistant asked each child, "Did the doctor touch you here?" Later in the interview, the child was asked to use the doll to show how the doctor had touched his or her buttocks and genitals.

Children were quite inaccurate across all conditions. Only 45% of the children who received genital examinations correctly answered "Yes" to the questions "Did the doctor touch you here [on buttocks or genitals]?" Only 50% of the children who did not receive genital exams correctly replied "No" to these questions. Further, the children's accuracy did not improve when they were given the dolls and asked to show how the doctor had touched them. Only 25% of the children who had received genital examinations correctly showed how the pediatrician had touched their genitals and buttocks. (A significant number of female subjects in this condition were inaccurate, because they inserted their fingers into the anal or genital cavities of the dolls—which the pediatrician never did.) Only 45% of the children who did not receive genital examinations were accurate in not showing any touching; that is, 55% of the children who did not receive genital examinations falsely showed either genital or anal touching when given the dolls, a pattern most prevalent among the females in this group; 75% of the females who did not receive a genital examination falsely showed that the pediatrician touched their genitals or their buttocks.

With the data on the potential usefulness of dolls equivocal at best, we feel that an important confound in the literature deserves mention: the context for the presentation of the dolls in these research settings is very different from that of actual forensic and clinical settings. Transcripts of therapy sessions with children suspected of having been sexually abused reveal interviewers employing various practices: naming the dolls after defendants; berating the dolls for alleged abuses against the child (e.g., shaking a finger at the male doll who has been named after the defendant and yelling, "You are naughty for hurting Jennifer!"); assuming the role of fantasy characters in doll play; and creating a persistent atmosphere of accusation. In the research settings in which the use of anatomical dolls has been studied, nonabused children were never subjected to such highly suggestive experiences prior to being interviewed with the dolls; they were not given prior motivation to play with the dolls suggestively or aggressively. On the other hand, children who were alleged to have been abused were sometimes exposed to the dolls repeatedly prior to coming to the

research setting; perhaps these interviews had involved repeated suggestions from parents and interviewers about various sexual themes. That their play with the dolls differed from that of nonabused children who lacked this prior experience could be attributed to the abused children's prior therapeutic or investigatory experiences, rather than to any inherent way in which they might be expected to play with the dolls.

Unfortunately, no study has examined the suggestive attributes of anatomical dolls, controlling for the preexperimental experience as a potentially serious confound. We simply do not know how nonabused children would behave with the dolls were they to have suggestive experiences prior to the experimental interview.[8] Conversely, we also do not know how abused children play with the dolls in their *first* investigatory interview, since the children in these studies have often been interviewed more than once and some have been exposed to the dolls at least once, prior to the experimental interview.

On the basis of our literature review (Ceci & Bruck, 1993), we concluded that the inconsistent findings point to the need for additional research and to the need for the development of explicit procedures to govern the use of anatomically correct dolls by interviewers. Until such research is available, the dolls ought to be used with great caution. Recently, Berry, & Skinner (1993) and Wolfner and his colleagues (1993) were even less supportive of doll use:

> . . . we are left with the conclusion that there is simply no scientific evidence available that would justify clinical or forensic diagnosis of abuse on the basis of the dolls. The common counter is that such play is "just one component" in reaching such a diagnosis based on a "full clinical" picture. . . . [Doll] play cannot be validly used as a component, however, unless it provides incremental validity, and there is virtually no evidence that it does." (Wolfner et al., p. 9)

Summary of current literature. The studies reviewed here highlight the different paradigms that researchers are now employing to examine children's suggestibility. In our review of this literature (Ceci & Bruck, 1993), we found that results of the most recent studies, in contrast to older ones, are somewhat more contradictory about the reliability of children's reports. One can locate studies claiming that young children are as immune to suggestion as older children (e.g., Marin, Holmes, Guth, & Kovac, 1979; Saywitz, et al., 1991), and studies claiming that younger children are more suggestible (Ceci, Ross, & Toglia, 1987; Cohen & Harnick, 1980; King & Yuille, 1987). Such mixed results have led to a confusing juxtaposition of headlines: "Study shows children are credible as witnesses." Or, "Research shows child witnesses unable to distinguish reality from fantasy."

A careful reading of the literature suggests, however, that there *are* reliable age differences in suggestibility, with preschoolers' reports more influenced by erroneous suggestions than older children's. In our review of the suggestibility literature, we found 18 studies that compared preschoolers to older children or to adults; in 15 of 18 of these studies, suggestibility was greater among preschoolers than older children or adults (see Table 2 in Ceci & Bruck, 1993). To be sure, some researchers attach various caveats to this conclusion. For example, some have claimed that age differences in suggestibility are evident mainly for nonparticipant children, i.e., bystanders (Rudy & Goodman, 1991); and for peripheral, nonsalient events (Fivush, 1993). And some researchers find that although young children may make some

errors in response to suggestive questions with a sexual theme, on the whole they are highly resistant to such questions (e.g., Saywitz et al., 1991; Goodman et al., 1991). Still others have found larger age differences in suggestibility for questions with sexual themes (e.g., Baker-Ward et al., in press; Goodman, Rudy, Bottoms, & Aman, 1990) and for questions about salient events (e.g., Cassel & Bjorklund, 1993).

Although preschoolers are usually depicted as being the most suggestible, it is important to point out that older children and adults are also suggestible. For example, as described above, 7-year-olds' reports, after 1 year, of their visits to the pediatrician could be quite easily altered through suggestion. Clarke-Stewart, Thompson, & Lepore (reported in Goodman & Clarke-Stewart, 1991) also found that 7-year-old children's reports and interpretations of a recently experienced event could be easily manipulated through suggestion. Also, Goodman, Wilson, Hazan, & Reed (1989) found that a substantial number of 7- to 10-year-old children incorrectly agreed with interviewers' suggestions about details of an event that occurred 4 years earlier. Many of these misleading suggestions had sexual themes. Finally, suggestions can alter some fundamental aspects of adults' autobiographical memories (Loftus, 1993). Thus, we cannot conclude that older children and adults are not suggestible, only that their level of suggestibility is less than that of preschoolers.

We reiterate, however, that the conditions created in these studies differ markedly from those that occur in actual therapy or in law enforcement investigations: these latter two contexts are seldom as sanitized of affect and free of motives as those in the research setting. The real life situation may entail high levels of stress, assaults to the child's body, and loss of control. In some cases, children are interviewed and reinterviewed under emotionally charged circumstances, entailing the use of bribes and threats, and often in the presence of highly distressed parents; under such conditions some children may finally utter reports that are simply consistent with the interviewer's expectations. In the McMartin case, interviewers were alleged to have coerced children's statements by praising them when they reported events that were consistent with the interviewer's beliefs and criticizing them for failing to do so (e.g., calling them "dumb"). Interviewers in both this case and other day-care cases also told children that other children had already disclosed the details of the abuse, thus creating added pressure to assent to suggestions of abuse. Not surprisingly, interviewers in the McMartin case managed to elicit statements of abuse from 369 of nearly 400 children they interviewed (Sauer, 1993), although only one child had made claims of abuse prior to the interviews. (This girl's accusations were so bizarre that the prosecution dropped them from the case [Sauer, 1993]).

Elsewhere we and others have used more emotionally laden events to examine issues related to the role of affect and bodily touching in producing misinformation effects, including suggestions about being kissed while naked, witnessing parents violate norms, or hurting others to protect loved ones (see Ceci, Leichtman, Putnick, & Nightingale, 1993), and experiencing painful and/or embarrassing medical procedures (e.g., Goodman, 1993; Ornstein, Baker-Ward, Gordon, & Merritt, 1993). Although children's resistance to suggestions are sensitive to all of these factors (and others), no study has attempted to incorporate all of them into a single experiment.

It is highly unlikely, however, that we will ever mimic the assaultive nature of some acts or interviews perpetrated on child victims and witnesses. Thus we are far from being able to provide a definitive conclusion about the reliability of all child witnesses' reports. It is safe to conclude,

though, that past pronouncements by some rather extreme advocates on both sides of the bench are simply unfounded. Children are neither as hypersuggestible and coachable as some prodefense advocates have alleged, nor as resistant to suggestions about their own bodies as some proprosecution advocates have claimed. They can be led, under certain conditions, to incorporate false suggestions into their accounts of even intimate bodily touching, but they can also be amazingly resistant to false suggestions and able to provide highly detailed and accurate reports of events that transpired weeks or months ago (e.g., Baker-Ward et al., in press). This mix of suggestibility and resistance to suggestion underscores the need for great caution in accepting the claims of those who would put either a prodefense or proprosecution "spin" on the data.

Policy Implications

Expert Witnesses

When a child comes to court to testify, this is often because he or she is the sole witness to a crime; this is particularly likely to be the situation in sexual abuse cases where the child is not only the sole witness, but there may be no physical evidence of abuse. The problem of uncorroborated testimony is compounded by the fact that the testimony of children may at times seem to lack credibility. As a result, both the prosecution and defense may call physicians, mental health professionals, and social scientists to serve as expert witnesses. In this section, we discuss the qualifications and roles of mental health professionals and social scientists who serve as expert witnesses in cases involving child witnesses, particularly in cases of alleged sexual abuse.

According to legal views (see Mason, 1991; Myers, 1993), these expert witnesses can be classified into two categories. (1) The first type, usually a mental health professional, is asked either to provide a generic description of the behavioral symptoms associated with sexual abuse, or to provide an opinion as to whether or not a particular child was abused. In the latter instance, the expert may have formulated his or her opinion based on therapy with the child or an assessment of a child's behavior; in some cases, the expert witness may have had no contact with the child in question. This type of expert is also frequently called upon to rehabilitate the credibility of a child witness who has been attacked by the defense, e.g., for delayed reporting. In this situation, the expert witness explains that, though such behaviors are *not* themselves diagnostic of abuse, it is not unusual for abused children to display a range of behaviors, such as recantation, delay of reporting, and inconsistent reporting. (2) The second type of expert witness is called to review the scientific literature on issues relevant to the credibility of child witnesses. This expert may cover various topics, including the literature on suggestibility as well as that on cognitive, emotional, and social development.

Although one might conclude that the research on children's suggestibility, discussed in this report, has policy implications for only the second type of witness, we argue that knowledge of this research is relevant to the professional qualifications and testimony of the first type of expert witness as well. The mental health professional who testifies on the diagnosis of sexual abuse or who describes to a court the symptoms associated with sexual abuse must also take into consideration competing hypotheses that might explain why the child in question, or children in general, demonstrate particular symptoms or make allegations of sexual abuse. One of the alternative hypotheses to be considered is that the particular child's allegations or symptoms have resulted from suggestive influences of the sort described above. It is important for the expert to consider

such an alternative, because those same symptoms associated with sexual abuse (delayed reporting, retraction of the allegation, inconsistent accounts, inappropriate knowledge of sexual behavior, or unusual play with anatomically correct dolls) have been observed in nonabused children who have been exposed to suggestive influences (see Berliner & Conte, 1993; Kendall-Tackett, Williams, & Finkelhor, 1993, for recent reviews of the literature).

Diagnosing child sexual abuse is thus a complex task requiring experience with sexually abused children and knowledge of both the clinical and the suggestibility and developmental literature. Experts who testify on such matters should be well-versed in these domains (Myers, 1993).

Some legal scholars have pointed out, however, that little experience and training is required of expert witnesses who provide testimony to rehabilitate the child witness. Myers (1993) writes that because this type of testimony is thought straightforward and simple, "a child protection services worker with six months on the job and knowledge of three or four pertinent articles is qualified to provide rehabilitative testimony on recantation and delayed reporting" (p. 177). We argue that this witness should have a more thorough knowledge of the scientific literature on both the indicators of child sexual abuse and the literature on suggestibility.

The second type of expert witness, those who testify about the scientific literature on suggestibility and child development, does not require clinical experience. However, this expert must have a thorough knowledge of the research literature germane to his or her testimony.

What the expert witness on children's suggestibility should tell the court. We come now to the question that has vexed any social scientist who ever dreamed (or had nightmares) of being called upon to serve as an expert witness or to prepare an amicus brief for an appellate court on children's testimony, namely, what does our present state of scientific knowledge permit us to say about the reliability of the testimony of the child witness? Having acknowledged the complexities of the research, we hold that expert witnesses, regardless of whether they are testifying for the prosecution or for the defense, should cover several points based on the literature:

1. There are reliable age effects in children's suggestibility, with preschoolers being more vulnerable than older children to a variety of factors that contribute to unreliable reports.

2. Although young children are often accurate reporters, some do make mistakes—particularly when they undergo suggestive interviews; and these errors can relate not only to peripheral details, but also to salient, predictable events that involve their own bodies.

3. Measures can be taken to lessen the risk of suggestibility effects. To date, the factors that we know most about concern the nature of the interview itself—its frequency, degree of suggestiveness, and demand characteristics.

- A child's report is less likely to be distorted, for example, after one interview than after several interviews (the term "interviews" here includes informal conversations between parents and child about the target events).
- Interviewers who ask nonleading questions, who do not have a confirmatory bias (i.e., an attachment to a single hypothesis), and who do not repeat close-ended, yes/no questions within or across interviews, are more likely to obtain accurate reports from children.

- Interviewers who are patient, nonjudgmental, and who do not attempt to create demand characteristics (e.g., by providing subtle rewards for certain responses) are likely to elicit the best quality reports from young children.

Thus, at one extreme we can have more confidence in a child's spontaneous statements made prior to any attempt by an adult to elicit what they suspect may be the truth. At the other extreme, we are more likely to be concerned when a child has made a statement after prolonged, repeated, suggestive interviews. Unfortunately, most cases lie between these extremes and require a case-by-case analysis.

4. Finally, it is also important that the court appreciate the complexity of the interrelationships of the factors affecting children's suggestibility. As in most areas of social science, effects are rarely as straightforward as one might wish. Even though suggestibility effects may be robust, the effects are not universal. Results vary between studies, and children's behavior varies within studies. Thus, even in studies with pronounced suggestibility effects, there are always some children who are highly resistant to suggestion. Some studies may show reliable age differences in suggestibility even though the majority of both younger and older children did not succumb to suggestion. We have seen this in our own studies as well as in transcripts of forensic and therapeutic interviews: in some cases, no matter how much an interviewer may try to suggest that an event occurred, some children will consistently resist and not incorporate the interviewer's suggestion or point of view. On the other side, although suggestibility effects tend to be most dramatic after prolonged and repeated interviewing, some children incorporate suggestions quickly, even after one short interview (e.g., Clarke-Stewart, et al., 1989, as reported in Goodman & Clarke-Stewart, 1991). No facile conclusion can be presented to courts on this matter.

Ideal vs. Actual Expert Witnesses. The "model" expert witness who comes forward to testify on issues related to children's suggestibility should be someone who has thoroughly reviewed the pertinent literature and who can present the relevant facts in a balanced manner to the triers of fact. This requirement is not an easy one to meet; this research area is developing rapidly and is riddled with a host of complex issues that necessitate a broad understanding of design, statistics, and theory not likely possessed by someone outside the research community.

Unfortunately, many who serve as expert witnesses do not have this breadth of knowledge. We have reviewed many examples of testimony by so-called experts that appeared to have been based on incomplete and at times dubious knowledge. Nowhere in their testimony is there any hint of the complexities that ought to have tamed the witness's statements to the jury. In the worst cases, the testimony was actually opposite to what we know to be the best evidence from systematic research. All too often such an expert appears in court strictly because his or her opinion is consistent with that of the defense or the prosecution, rather than because the witness is truly knowledgeable about the field. Such testimony can be a disservice to the aims of justice, not to mention to the professions these expert witnesses represent.

Although the above discussion pertains mainly to the social scientist who testifies about children's suggestibility, our review of case material and the literature suggests that these same criticisms can be made of some experts who testify about the behavioral symptoms associated with sexual abuse. Mason (1991) analyzed 122 civil and criminal appellate court cases in which expert witnesses testified about child sexual abuse. She

found that experts frequently presented testimony that was either internally inconsistent or was contradicted by other experts. For example, 14 experts cited age-inappropriate knowledge of sex and sexual preoccupations as characteristics of an abused child, whereas 6 experts asserted that naivete and aversion to sexual matters characterized the sexually abused child. Some experts maintained that consistent accounts of events were important indicators of sexual abuse, whereas others maintained the opposite, that sexually abused children are characterized by their inconsistent accounts. Mason also reported that appellate courts tend to take expert testimony at face value; that they rarely raise questions about the testimony's acceptance by the scientific community, or about the credentials of the mental health professional presenting expert testimony.

The response of the courts in this study reveals that there is a critical gulf between the scientific community and the judiciary. Judges are not willing and probably not able to critically evaluate the reliability of the testimony offered (Mason, 1991, p. 205).

An example of the unscientific nature of some experts' testimony is illustrated by *Kelly Michaels* v. *State of New Jersey*. Michaels was a preschool teacher convicted on 115 counts of sexual offenses involving 20 children, and sentenced to 47 years in prison. The expert witness for the prosecution testified that conduct of all but one of the child witnesses was consistent with having been sexually abused. She did not seriously consider the possible effects of numerous suggestive influences on the children's testimony, which had included persistent, aggressive, and suggestive interviews with children who initially denied that anything had happened. In this case, the appellate court did challenge the testimony of the expert witness, reversing Michael's conviction (after she had spent 5 years in prison), in part because the expert's testimony concerning the child behavioral indicators of abuse did not have acceptance within the relevant scientific community.

Problems with expert testimony are endemic to our legal system and to those of other countries whose codes have been derived from common law. As far back as one can check, jurists and laypersons alike have viewed expert witnesses as untrustworthy, as inclined to put a "spin" on interpretations of the data toward the side that hired them:

> Perhaps the testimony which least deserves credit with a jury is that of skilled witnesses. . . . It is often quite surprising to see with what facility, and to what extent, their views can be made to correspond with the wishes and interests of the parties who call them. (Judge John Pitt Taylor, 1858, p. 65-69, as quoted in Gross, 1991).

These views continue to be expressed by American jurists:

> To put it bluntly, in many professions service as an expert witness is not generally considered honest work. . . . Experts in fields see lawyers as unprincipled manipulators of their disciplines, and lawyers and experts alike see expert witnesses—those members of the learned professions who will consort with lawyers—as whores. (Gross, 1991, p. 1115)

And by British jurists:

> Expert evidence is sometimes given by people whose level of knowledge seems lamentably low. A number of the recent, and best-known scandals show this. . . . How does this come about? In the

> first place, I think it is because our present system provides no systematic quality control. Broadly speaking, anyone can be an expert witness, provided they have some relevant knowledge, and nothing whatever is done to see that only the best people are used. To be allowed to give expert evidence, witnesses must satisfy the judge that they have some practical experience, or some professional qualifications; but that is all. No minimum standards are laid down. The only test is opposing counsel's cross-examination; and, in a jury trial, this may be designed to score clever points, rather than to test whether they (i.e., the experts) are really good at their job. (Spencer, 1992, pp. 216-217)

The Relationship of Research to Clinical Practice

That the judicial community is unwilling or unable to evaluate critically the testimony of social science experts and mental health professionals reflects to some degree an incomplete or inaccurate understanding of the relevant knowledge base, but it also reflects, in the case of the reliability of children's reports, a gulf between clinical practice and social science research. As a result, in those cases where clinical practice is not informed by research findings, clinicians and social scientists may present diametrically opposite expert testimony on the very same topic. Two examples illustrate this breach between the two perspectives.

The first involves a survey of 212 mental health professionals about their assessment and validation procedures in sexual abuse cases (Conte et al., 1991). Of relevance to the present report, it was found that children had already been asked to tell their story an average of 2.3 times before talking to the professional respondent; only 27% of respondents indicated that they were the first person to talk with the child about the abuse. In discussing these findings, however, the authors do not seriously consider the impact of such interviewing practices:

> Little is currently known about the effects of such prior interviewing on the child's willingness to engage with yet another adult or on the quality of information obtained from the child. While some professionals are likely to make much of the possible "contamination" that these prior interviewers have on the child's reports, there are virtually no data currently available suggesting that adults have the power through interviewing techniques to alter fundamentally a child's understanding of and ability to describe what events did or did not take place. (p. 433).

We hope that this report will begin to inform professionals that such data are available.

A second example that illustrates the gulf between practice and research concerns the use of anatomically correct dolls. Many professionals have no formal training or experience in the use of the dolls (Boat & Everson, 1988) and may view some interactions of children with the dolls (e.g., placing a finger in the doll's anal cavity, tugging on its penis, or avoiding the dolls altogether) as indicative of sexual abuse, even though there is no scientific support that such interactions are diagnostic of abuse. In a recent survey, for example, only 16% of mental health and law professionals stated that avoidance of the dolls was normal,

while 80% rated digital penetration as abnormal (Kendall-Tackett, 1991). Yet, as reviewed above, such behaviors are commonly observed in nonabused children.

Of more concern, perhaps, is the American Psychological Association's (1991) current position on the use of the dolls. The following statement was issued by APA's Council of Representatives:

> Neither the dolls nor their use are standardized or accompanied by normative data. . . . We urge continued research in quest of more and better data regarding the stimulus properties of such dolls and normative behavior of abused and nonabused children. . . . Nevertheless, doll-centered assessment of children, when used as part of a psychological evaluation and interpreted by experienced and competent examiners, may be the best available practical solution for a pressing and frequent clinical problem. (APA, 1991, p. 1).

The APA's policy position seems contradictory in its noting first that there are no standardized methods for doll interviews or normative data on nonabused or abused children's doll play, but then asserting that experienced interviewers may nevertheless find doll-centered assessment the best available method for evaluating children suspected of having been sexually abuse. Even if one assumes that experienced examiners can avoid making false inferences from children's doll play, and that such doll play can provide important clinical insights not obtainable from other sources, the APA should nevertheless codify this expert knowledge in such a way that researchers can accurately assess the incremental validity of doll-based assessments. Our reading of the literature is that at present such knowledge is more illusory than real (see Wolfner et al.'s criticism [1993] of the lack of incremental validity of doll-based assessments). Even if anatomical dolls are used as just one part of an assessment, other aspects of so-called "developmentally sensitive assessments" (e.g., play therapy, role playing, techniques that induce visually-guided imagery, self-empowerment training) may interact with the doll use to produce false positive assertions of abuse. Because the appropriate research has yet to be done, it is shortsighted to assume (as some experts have testified in court) that the dolls do not present reliability risks. Although it could be the case that the use of dolls does provide important information, it could also be the case that this method leads to unacceptable levels of false positive reports. Only research will tell.

The fact remains that clinicians and mental health professionals face many dilemmas and choices in providing for children who may have been sexually abused. Often the favored choice may conflict with forensic procedures. Let us consider one scenario: a child has been removed from her home as a result of a report of sexual abuse and has been placed in emergency foster care, separated from her family, friends, and school. The child is greatly distressed and in need of immediate counseling. The forensic interviews will not be completed for several months. In light of some research findings that children's reports are likely to be more accurate if interviews (which include therapy sessions) concerning the alleged abuse are held to a minimum until after the forensic interview takes place, when should the mental health professional begin therapy with the child? How can we avoid the twin dangers of, on the one hand, putting the child's emotional needs on hold until after the forensic interviews and, on the other hand, providing counseling that can be potentially damaging to the veracity of the child's report? We know of no easy answers.

Given the pressing needs of both sides in a criminal dispute to prepare, investigate, and often reinterview, no amount of child-friendly court procedures can totally alleviate some of the problems associated with children's testimony. Yet, perhaps there are ways of providing therapeutic support that lessen the likelihood of tainting the child's report. Therapeutic procedures that involve visually-guided imagery in the context of the abuse-related allegations might be avoided, as might forms of therapy that make contact with the abusive scenario (e.g., self-empowerment training, role playing, doll use, hypnosis).

Although some might argue that it would be too restrictive and ultimately damaging to a child's development were therapists to avoid potentially suggestive techniques, it could also be argued that employing such interventions simply constitutes too great a risk. On the one hand, if the defendant is innocent, such techniques could promote and reinforce false allegations. On the other hand, if the defendant is guilty, these interventions may end up discrediting the child's testimony, with defense attorneys arguing that the child's reports are the product of highly suggestive therapeutic techniques. Finally, on the empirical side, we are unaware of any persuasive treatment–outcome validity research indicating that suggestive techniques are necessary in therapy to achieve a positive mental health outcome for children suspected of being abused. Given this state of knowledge, clinicians might consider limiting interventions to *non*suggestive techniques in therapy until young clients have given sworn statements; such an approach may afford minimal danger to the child.

Professional Organizations as Ethical Gatekeepers

Professional organizations could help resolve some of the problems we have been discussing by making ethics codes for expert witnesses more explicit. Existing codes for expert witnesses of the organizations that represent various constituencies (psychology, social work, pediatrics) tend to be weak and ill-defined—in part because "expert witness" is an ill-defined legal concept.[9] The Federal Rules of Evidence 702 states that if scientific, technical, or other specialized knowledge will assist a fact finder in understanding evidence, then a witness may be regarded as an expert by virtue of his or her knowledge, skill, experience, training, or education. This rule construes expertise broadly enough to cover all fields, including emerging areas within fields, and is constrained by two other Federal Rules of Evidence (401 and 403), which specify that the expert testimony must be relevant. Together, these Federal Rules allow virtually anyone who possesses an advanced degree, or who has some clinical experience, to offer expert testimony on children's credibility, even though the expert may have scant knowledge of the current scientific findings. As a result, experts testifying in child sexual abuse cases have offered totally opposite interpretations of children's behavior and testimony (see Mason [1991] above).

To some extent, weak ethics codes also reflect the ascendancy of guild interests. Because no constituency wants to be excluded from activities that involve service to others (at times for financial gain), its representatives ensure that its members' role is not diminished by ethics code language. Ethics codes tend to be explicit about matters that are relatively benign to the group as a whole (e.g., rules for preparing reports, or statements regarding generic conflicts of interest), but vague about matters that could adversely affect the entire membership (e.g., defining precisely what an expert should know in order to testify about children's suggestibility, or what it means to conduct a good interview). Thus, for example, when psychologists look to their own specialty guidelines and general ethics codes for guidance about the credentials or conduct of an expert witness, they find

little help other than enjoinders to act responsibly, to be informed, and to aspire to the norms that guide a professional toward the highest ideals. Consider some of the sections of the most recent APA code of ethics revision (Ethical Principles, 1992) relevant to forensic services:

- Psychologists appropriately take into account the ways in which a prior relationship might affect their professional objectivity or opinions and disclose potential conflict to the parties. (Section 705)
- Psychologists who engage in . . . professional activities maintain a reasonable level of awareness of scientific and professional information in their fields. (General Standards 1.5)
- Psychologists rely on scientifically and professionally derived knowledge when making scientific or professional judgments. (General Standards 1.6)
- In addition, psychologists base their forensic work on appropriate knowledge of a competence in the areas underlying such work. (General Standards 7.1)
- In forensic testimony and reports, psychologists . . . describe fairly the bases for their testimony and conclusions [and] whenever necessary to avoid misleading, acknowledge the limits of their data or conclusions. (General Standards 7.4).

However well intended, these statements taken together lend themselves to ambiguous interpretation. For example, can a psychologist, in testifying about children's suggestibility, rely on *either* research knowledge *or* clinical experience? Can a therapist be expected to avoid a conflict of interest and maintain sufficient objectivity to serve as an expert witness when he or she has had extended contact with the child? (Apparently so. Mason's analysis [1991] showed that many expert witnesses who testified in abuse cases were often the child's therapist, and only 13% of all experts had no prior relationship with the child.)

Missing from ethics codes and specialty guidelines for expert witnesses (e.g., Committee on Ethical Guidelines, 1991) is language that would specify that they bring to court more than an advanced degree, a supervised internship which had brought them into contact with sexually abused children, or other clinical experience whereby they had occasionally seen sexually abused clients in their practice. An expert testifying on children's suggestibility, and more generally on the credibility of child witnesses, should be intimately familiar with the systematic scholarship on the topic. Although it is not necessary for this expert to be a researcher, he or she needs to be at least a critical consumer of the research literature.

The failure of professional organizations to constitute and then to enforce principled guidelines has serious consequences. First, it can undermine the judicial system's confidence in the capacity of professionals to offer reliable testimony. More important, in criminal proceedings where the defendant faces incarceration, or in civil proceedings where the future placement of the child is at stake, the legitimacy of the expert's testimony can be critical to preserving the rights of both the child and the defendant.

To conclude, enforcement mechanisms are needed to ensure that expert testimony can be evaluated for its scientific merit. Until such mechanisms are openly advertised to all consumers of legal services, enjoinders to "stay informed" will probably do little to ebb the sorts of abuses reported

by Mason (1991) and Spencer (1992). Because of their vagueness, professional ethics codes will be implemented more often in the breach than in the letter.

Conclusion

We have argued that the investigation of child sexual abuse allegations and expert testimony addressing such investigations are fraught with problems. Scientists have begun to contribute important insights to these problems, though clearly more research is needed. We have provided some troubling examples of how research has failed to inform practice, and how experts often go beyond what current scientific findings seem to warrant.

To be sure, those charged with investigating, reporting, and treating suspected child maltreatment face immense obstacles. These professionals are deeply aware of the pervasiveness of child sexual abuse, and the all too frequent ineffectiveness of prosecution. They know better than most the emotionally wrenching sequelae of abuse, especially intrafamilial abuse. And they are keenly aware that the evidence from research must always be tempered by real-world considerations, no matter what the "significance level" or "effect size" of a finding. Thus, many pressing, unresolved issues concerning the interviewing and treatment of individual children remain.

We presented a scenario highlighting the difficult task faced by many professionals on a daily basis—how to promote two goals that often conflict: how to provide for the child's mental health needs while simultaneously protecting the legal rights of the accused. While we must strive to uncover abuse, we must eschew interview processes that may promote false beliefs, fantasies, or fabrications—regardless of the nature of the initiating event. Just as we have argued above that it is unethical for social scientists to institute experimental manipulations that might change the fundamental nature of children's emotionally salient autobiographical memories, it is equally indefensible for therapists or forensic interviewers to cause such changes. The results of persistent erroneous suggestions and of failures to test alternative hypotheses can be lasting, as evidenced by the experiences and reactions of the child witnesses in the McMartin trials, described at the beginning of this report:

> No one who saw them will soon forget the frenzied faces of . . . former McMartin pupils (who) had spent their last six years—fully half their lives—instructed in the faith that they had been subjected, at ages 4 and 5, to unspeakable sexual horrors; this belief they had come to hold as the defining truth of their lives and identities. It is not surprising that these children should have wept and raved when the verdict was handed down denying all that they believed in. (Rabinowitz, 1990, p. 63)

Notes

[1] In view of the public outcry against the seeming refusal by jurors to believe the children, posttrial statements by these same jurors about believing some of the children's claims may have been self-serving. One close observer of the trial suggested this possibility to us.

[2] Several have been the focus of books. The Wee Care case involving Kelly Michaels was the source for several books, including *Naptime* (Manshel, 1990) and *Not My Child* (Crowley, 1990). Other cases have been detailed in television documentaries (e.g., the Little Rascals case involving Bob Kelly and five other defendants, which was the focus of three Frontline documentaries, e.g., *Loss of Innocence*); movies (e.g., the Country Walk day-care case in Miami, which was the basis of the movie *Unspeakable Acts*); and

magazine and newspaper articles (e.g., Nathan, 1987; Rabinowitz, 1990).

[3]Although this report focuses on the interviewing of alleged child sexual abuse victims, the literature reviewed is equally important to nonabuse cases that involve the child witness. Sexual abuse is of special interest, because this category of complaints appears to represent the single largest class of actions that eventuate in criminal court testimony (as opposed to neglect cases or custody disputes which are largely litigated in juvenile and family court systems). Our discussion centers on nursery school cases, because, although these cases represent only a small proportion of sexual abuse complaints, in absolute numbers they involve a large number of children (in the McMartin case, for instance, interviewers under contract to the State of California alleged the abuse of 369 children [Sauer, 1993]); moreover, day-care cases are relevant to the more general testimonial issues found in many nonday-care cases (i.e., repeated suggestive questioning, interviewer stereotypes, failure to test alternative hypotheses). Finally, because of their visibility, day-care cases are often more extensively documented.

[4]Cases are classified as *substantiated* or *indicated* based on how consistent the evidence from an investigation is with abuse; in most states this is a matter of caseworker judgment. States usually have a two-tiered system of classifying investigations as either substantiated/founded, on the one hand, or unsubstantiated/not founded, on the other. Some states use a third tier that is intermediate between substantiated and unsubstantiated, namely, "indicated." This term is given to cases in which the agency doing the investigation may have "reason to suspect" that abuse occurred, but the level of evidence does not rise to the level required for the designation "substantiated." The lowest level of evidence needed to substantiate a case is "some credible evidence," which is used by 18 states, while the highest level of evidence needed is "preponderance of evidence," which is used by 12 states. An intermediate level of evidence is used by an additional 12 states, and the remaining states use idiosyncratic terminology (see Figure 5 of National Center of Child Abuse and Neglect, 1993, p. 28). Thus, lower levels of evidence increase the possibility that, upon further investigation, a subsequent determination may be made that insufficient evidence exists to designate the presence of abuse. For these reasons it is important not to use the terms "substantiation," "indication," and "validation" interchangeably.

[5]Some have suggested that these two figures be summed to 8%. This assumes, however, that there were different children in the two categories, which is not clear from the published report. The breakdown reported here is that reported by Saywitz, et al. (1990).

[6]Courts have taken notice of the need to distinguish between an interviewer whose view reflects a strongly held expectation versus the interviewer who possesses relevant background information. For example, in *Idaho* v. *Wright* the Court accepted the argument contained in an amicus brief that "there is an important distinction between preconceptions that can cloud judgment, and background information that is needed for thorough evaluation of sexual abuse" (Amicus Brief to the Supreme Court in *Idaho* v. *Wright*, No. 89-260, p. 96).

[7]These data reveal an interesting disjunction with the reasoning that when children retract earlier claims of sexual abuse, this is indicative, if not diagnostic, of a truthful original report (Sgroi, 1982; Summit, 1983). In this study, it was often the case that children originally made false allegations, which they then, with gentle persuasion, recanted. Were this finding applicable to situations that are abuse-related--and we make no such claim here—it could be suggested that retraction might also be consistent with an erroneous original report.

[8]Pilot data from one subject addresses this question. A 3½-year-old nonabused girl was examined by a pediatrician. She was not given a genital examination. Immediately after the examination, when interviewed by the experimenter, she correctly said that the doctor had not touched her genitals or buttocks. Furthermore, when shown an anatomically correct doll and told to show how the doctor had touched her genitals and buttocks, she correctly stated that he had not touched her. Three days later, the same child was shown the anatomically correct doll and asked to show all the things that the doctor did to her in her previous visit. This time, she inserted a stick into the vagina of the doll. Upon further questioning, however, she said that the doctor did not do this. Three more days later, the child was asked to use the anatomically correct doll to show her father everything that had happened at the examination. This time, she hammered a stick into the doll's vagina and then inserted a toy earscope into the doll's anus. When asked if this really happened, she said "Yes it did." When her father and the experimenter both tried to debrief her with such statements as, "Your doctor doesn't do those things to little girls. You were just fooling. We know he didn't do those things," the subject clung tenaciously to her claims. Thus, for this one subject, repeated exposure to the doll, with minimal suggestions, resulted in highly sexualized play. It is critical that such a finding be replicated with a large, diverse sample to determine if this child's response is representative of nonabused children.

[9]For example, most courts disallow expert testimony that speaks directly to the *ultimate question*, that is, the defendant's

guilt or innocence. In some courts, however, expert witnesses are permitted to testify as to whether they believe the child was abused (see Myers, 1992). One would think that an expert's opinion that a particular child was abused might have the same effect as speaking to the child's credibility. This leads to confusion even among the legal scholars whom we have consulted, with one remarking that the courts' thinking regarding this issue is little more than "wordplay."

References

American Psychological Association. (1991). *Statement on the use of anatomically detailed dolls in forensic evaluations.* Washington, DC: APA Council of Representatives.

Andrews, J. A. (1964). The evidence of children. *Criminal Law Review, 64,* 769-777.

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

Batterman-Faunce, J. M., & Goodman, G. S. (1993). Effects of context on the accuracy and suggestibility of child witnesses. In G. S. Goodman & B. L. Bottoms (Eds.), *Child victims and child witnesses: Understanding and improving testimony* (pp. 301-330). New York: Guilford Press.

Berliner, L., & Conte, J. R. (1993). Sexual abuse evaluations: Conceptual and empirical obstacles. *Child Abuse and Neglect, 17,* 111-125.

Berry, K., & Skinner, L. G. (1993). Anatomically detailed dolls and the evaluations of child sexual abuse allegations: Psychometric considerations. *Law and Human Behavior, 17,* 399-422.

Besharov, D. (1991). Child abuse and neglect reporting and investigation: Policy guidelines for decision making. *Child and Youth Services, 15,* 35-50.

Binet, A. (1900). *La suggestibilité.* Paris: Schleicher Frères.

Boat, B., & Everson, M. (1988). The use of anatomical dolls among professionals in sexual abuse evaluations. *Child Abuse and Neglect, 12,* 171-186.

Brown, M. R. (1926). *Legal psychology.* Indianapolis: Bobbs-Merrill.

Bruck, M., Ceci, S. J., Francoeur, E., & Barr, R. (in press). *"I hardly cried when I got my shot": Young children's reports of their visit to a pediatrician.* Manuscript submitted for puolication.

Cassel, W. S., & Bjorklund, D. F. (1993). *Tell me about . . . . Don't you remember . . . ? Isn't it true that . . . ? Developmental patterns of eyewitness responses to increasingly suggestive questions.* Unpublished manuscript, Florida Atlantic University, Boca Raton, FL.

Ceci, S. J. (in press). Cognitive and social factors in children's testimony. In B. Sales & G. VandenBos (eds.), *APA Master Lectures: Psychology and the Law.* Washington, DC.: American Psychological Association.

Ceci, S. J., & Bronfenbrenner, U. (1991). On the demise of everyday memory: The rumors of my death are greatly exaggerated. *American Psychologist, 46,* 27-31.

Ceci, S. J. & Bruck, M. (1993). The suggestibility of the child witness. *Psychological Bulletin, 113,* 403-439

Ceci, S. J., Leichtman, M., Putnick, M., & Nightingale, N. (1993). Age differences in suggestibility (pp. 117-138). In D. Cicchetti & S. Toth (Eds.), *Child witnesses, child abuse, and public policy.* Norwood, NJ: Ablex.

Ceci, S. J., Leichtman, M. D., & Bruck, M. (In press). The suggestibility of children's eyewitness reports: Methodological issues. In F. Weinert & W. Schneider (Eds.), *Memory development: State of the art and future directions* . Englewood Cliffs, NJ: Erlbaum.

Ceci, S. J., Leichtman, M., & White, T. (in press). Interviewing preschoolers: Remembrance of things planted. In D. P. Peters (Ed.), *The child witness in context: Cognitive, social, and legal perspectives.* Holland: Kluwer.

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

Chadbourn, J. (1978). *Wigmore on evidence.* Boston: Little, Brown.

Charlier, T., & Downing, S. (1988, January 17-22). Justice abused: A 1980's witch hunt. Memphis *Commercial Appeal.*

Clarke-Stewart, A., Thompson, W., & Lepore, S. (1989, April). *Manipulating children's interpretations through interrogation.* Paper presented at the biennial meeting of the Society for Research in Child Development, Kansas City.

Cohen, R. L. (1975). Children's testimony and hearsay evidence. *Law Reform Commission of Canada. Report on evidence.* Ottawa: Information Canada.

Cohen, R. L., & Harnick, M. A. (1980). The susceptibility of child witnesses to suggestion. *Law and Human Behavior, 4,* 201-210.

Committee on Ethical Guidelines for Forensic Psychologists. (1991). Specialty guidelines for forensic psychologists. *Law & Human Behavior, 15,* 655-665.

Conte, J. R., Sorenson, E., Fogarty, L., & Rosa, J. D. (1991). Evaluating children's reports of sexual abuse: Results from a survey of professionals. *Journal of Orthopsychiatry, 78,* 428-432.

Cov v. *Iowa,* 108 S. Ct. 2798 (1988).

Crowley, P. (1990). *Not my child: A mother confronts her*

child's sexual abuse. New York: Doubleday.

Daro, D., & Mitchel, L. (1990). *Current trends in child abuse reporting and fatalities: The results of the 1989 annual 50 states survey*. Washington, DC: National Commission for the Prevention of Child Abuse.

Dawes, R. (1992, Spring). The importance of alternative hypothesis and hypothetical counterfactuals in general social science. *The General Psychologist*, 2-7.

Doris, J. (1993). Child witness conference. Supplemental RFP Children's Justice and Assistance Act funds. Submitted by the Family Life Development Center, Cornell University.

Ethical principles of psychologists and code of conduct. (1992). *American Psychologist, 47*, 1597-1611.

Federal Rules of Evidence, Rule 702.

Fischer, M. A. (1988). Flip-flop: Why four years later the press is taking a strikingly different approach to the McMartin preschool scandal. *Los Angeles Magazine, 33*, 85-94.

Fivush, R. (1993). Developmental perspectives on autobiographical recall. In G. S. Goodman & B. L. Bottoms (Eds.), *Child victims and child witnesses: Understanding and improving testimony* (pp. 1-24). New York: Guilford Press.

Flin, R. (1993). Hearing and testing children's evidence. In G. S. Goodman & B. L. Bottoms (Eds.), *Child victims and child witnesses: Understanding and improving testimony* (pp. 279-300). New York: Guilford Press.

Gardner, R. (1989). *Sex abuse hysteria: Salem witch trials revisited*. Longwood, NJ: Creative Therapeutics Press.

Glaser, D., & Collins, C. (1989). The response of young, non-sexually abused children to anatomically correct dolls. *Journal of Child Psychology and Psychiatry, 30*, 547-560.

Goodman, G. S. (1993, March). *Children's memory for stressful events: Theoretical and developmental considerations*. Paper presented at the biennial meeting of the Society for Research in Child Development. New Orleans.

Goodman, G. S., & Aman, C. (1990). Children's use of anatomically detailed dolls to recount an event. *Child Development, 61*, 1859-1871.

Goodman, G. S. & Clarke-Stewart, A. (1991). Suggestibility in children's testimony: Implications for child sexual abuse investigations. In J. L. Doris (Ed.), *The suggestibility of children's recollections* (pp. 92-105). Washington, DC: American Psychological Association.

Goodman, G. S., Hirschman, J. E., Hepps, D., & Rudy, L. (1991). Children's memory for stressful events. *Merrill Palmer Quarterly, 37*, 109-158.

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

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.

Gray, E. (1993). *Unequal justice: The prosecution of child sexual abuse*. New York: MacMillan.

Gross, S. R. (1991). Expert evidence. *Wisconsin Law Review, 1113*, 1114-1231.

Harvard Law Review Notes. (1985). The testimony of child sex abuse victims in sex abuse prosecutions: Two legislative innovations. *Harvard Law Review, 98*, 806-827.

Idaho v. *Wright*, 110 S. Ct. 3139 (1990).

Kendall-Tackett, K. (1991, April). *Professionals' standards of "normal" behavior with anatomical dolls and factors that influence these standards*. Paper presented at the biennial meeting of the Society for Research in Child Development, Seattle.

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.

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

Leichtman, M. D., & Ceci, S. J. (revision under review). The effect of stereotypes and suggestions on preschoolers' reports. *Developmental Psychology*.

Lipmann, O. (1911). Pedagogical psychology of report. *Journal of Educational Psychology, 2*, 253-261.

Loftus, E. F. (1993). The reality of repressed memories. *American Psychologist, 48*, 518-537.

Manshel, L. (1990). *Nap time*. New York: Kensington.

Marin, B. V., Holmes, D. L., Guth, M., & Kovac, P. (1979). The potential of children as eyewitnesses. *Law and Human Behavior, 3*, 295-305.

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

Mason, M. A. (1991, Winter-Fall). A judicial dilemma: expert witness testimony in child sex abuse cases. *Psychiatry and Law*, 185-219.

McGough, L. (in press). *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(Serial No. 229), 153-159.

Messerschmidt, R. (1933). The suggestibility of boys and girls between the ages of six and sixteen. *Journal of Genetic Psychology, 43*, 422-437.

Montoya, J. (1992). On truth and shielding in child abuse trials. *Hastings Law Journal, 43*, 1259-1319.

Montoya, J. (1993). *When fragile voices intersect with a fragile process: Pretrial interrogation of child witnesses.* Unpublished monograph. University of San Diego School of Law.

Moss, D. C. (1988). "Real" dolls too suggestive: Do anatomically correct dolls lead to false abuse charges? *American Bar Association Journal, 24*, 24-25.

Myers, J. E. (1992). *Legal issues in child abuse and neglect.* Newbury Park, CA: Sage.

Myers, J. E. (1993). Expert testimony regarding child sexual abuse. *Child Abuse and Neglect, 17*, 175-185.

Nathan, D. (1987). The making of a modern witch trial. *The Village Voice, 33, 19-32.*

National Center for Child Abuse and Neglect. (1988). *Study of national incidence and prevalence of child abuse and neglect: 1988.* Gaithersburg, MD: U.S. Department of Health and Human Services.

National Center for Child Abuse and Neglect. (1993). *National Child Abuse and Neglect Data System. 1991: Summary data component.* Gaithersburg, MD: U.S. Department of Health and Human Services.

New Jersey v. *Michaels* (1993). 264 N.J. super (AppDiv) 579.

Oates, K., & Shrimpton, S. (1991). Children's memories for stressful and nonstressful events. *Medicine, Science, and the Law, 31*, 4-10.

Ornstein, P. A., Baker-Ward, L., Gordon, B., & Merritt, K. (1993). *Children's memory for medical procedures.* Paper presented at the biennial meeting of the Society for Research in Child Development, New Orleans.

Ornstein, P. A., Gordon, B. N., & Larus, D. (1992). Children's memory for a personally experienced event: Implications for testimony. *Applied Cognitive Psychology, 6*, 49-60.

Otis, M. (1924). A study of suggestibility in children. *Archives of Psychology, 11*, 5-108.

Peters, D. D., Wyatt, G. E., & Finkelhor, D. (1986). Prevalence. In D. Finkelhor (Ed.), *A sourcebook on child sexual abuse.* Beverly Hills: Sage.

Popper, K. R. (1962). *Conjectures and reflections.* New York: Basic Books.

Rabinowitz, D. (1990, May). From the mouths of babes to a jail cell: Child abuse and the abuse of justice. *Harper's Magazine*, pp. 52-63.

Raskin, D., & Esplin, P. (1991). Assessment of children's statements of sexual abuse. In J. L. Doris (Ed.), *The suggestibility of children's recollections* (pp. 153-164). Washington, DC: American Psychological Association.

Raskin, D. & Yuille, J. (1989). Problems in evaluating interviews of children in sexual abuse cases. In S. J. Ceci, M. P. Toglia, & D. F. Ross (Eds.), *Adults' perceptions of children's testimony* (pp. 184-207). New York: Springer-Verlag.

Robin, M. (1991). The social construction of child abuse and "false allegations." *Child and Youth Services, 15*, 1-34.

Rudy, L., & Goodman, G. S. (1991). Effects of participation on children's reports: Implications for children's testimony. *Developmental Psychology, 27*, 527-538.

Sauer, M. (1993, August 29). Decade of accusations. *San Diego Union Tribune*, D1-D3.

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

Schooler, J. W., Gerhard, D., & Loftus, E. F. (1986). Qualities of the unreal. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 12*, 171-181.

Sgroi, S. M. (1982). *Handbook of clinical intervention in child sexual abuse.* Lexington, MA: Lexington Books.

Shaw, D. (1990, January 19-22, Series). McMartin verdict: Not guilty. *Los Angeles Times.*

Sherman, I. (1925). The suggestibility of normal and mentally defective children. *Comparative Psychology Monographs, 2.*

Siegel, J. M., Sorenson, S. B., Golding, J. M., Burnam, M. A., & Stein, J. A. (1987). The prevalence of childhood sexual assault: The Los Angeles epidemiologic catchment area project. *American Journal of Epidemiology, 126*, 1141-1153.

Spencer, J. R. (1992). Court experts and expert witnesses. In R. V Rideout & B. Hepple (Eds.), *Current legal problems, Vol. 45* (pp. 214-236). London: Oxford University Press.

Steller, M. (1991). Commentary: Rehabilitation of the child witness. In J. L. Doris (Ed.), *The suggestibility of children's recollections* (pp. 106-109). Washington, DC: American Psychological Association.

Stern, W. (1910). Abstracts of lectures on the psychology of testimony and on the study of individuality. *American Journal of Psychology, 21*, 270-282.

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

Taylor, J. P. (cited in S. R. Gross) (1858). *Treatise on the law of evidence, 3rd Ed.* §§ 45-50, 65-69.

Terr, L. (1988). Anatomically correct dolls: Should they be used as a basis for expert testimony? *Journal of the American Academy of Child and Adolescent Psychiatry, 27*, 254-257.

Varendonck, J. (1911). Les témoignages d'enfants dans un proces retentissant [The testimony of children in a famous trial]. *Archives de psychologie, 11*, 129-171.

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.

Wyatt, G.E. (1985). The sexual abuse of Afro-American and white American women in childhood. *Child Abuse and Neglect, 9*, 507-519.

About the Authors

Steve Ceci, Ph.D., is Helen L. Carr Professor of Developmental Psychology in the Department of Human Development and Family Studies, Cornell University.

Maggie Bruck, Ph.D., is an Associate Professor in the departments of Psychology and Pediatrics at McGill University.

Acknowledgements

This research was supported, in part, by grants to the first author from the National Institute of Child Health and Human Development and to the second author by a grant from the Natural Sciences and Engineering Research Council.

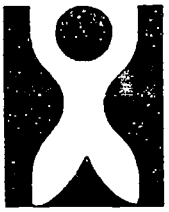
Next issue: (Winter 1993). Integrating science and ethics in research with high-risk and underserved children and youth [tentative title]. Celia B. Fisher.

Subscriptions ($10 to nonmembers of SRCD) and single issues ($3.50) are available from:

SRCD Executive Office
University of Michigan
300 North Ingalls, 10th floor
Ann Arbor, MI 48109-0406

Social Policy Report *is a quarterly publication of the Committee on Child Development, Public Policy and Public Information of the Society for Research in Child Development.*

Editor: Nancy G. Thomas
Phone: (303) 925-5516
Fax: (303) 925-9570



SRCD Executive Office
University of Michigan
300 N. Ingalls 10th Floor
Ann Arbor, MI 48109-0406

SOCIAL POLICY REPORT