TABLE OF CONTENTS

Table of Authorities

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

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

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

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

Bruck, M., Ceci, S.J., Melnyk, L., & Finkelberg, D. (1999). The effect of interviewer bias on the accuracy of children's reports and interviewer's reports. Paper presented at the Biennial Meeting of the Society for Child Development. Albequerque, NM. .

Bruck, M., Melnyk, L., & Ceci, S.J. (2000). Draw it again Sam: The effects of repeated drawing on the accuracy of children's reports and source monitoring attributions. *Journal of Experimental Child Psychology*. Accepted for publication

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

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

Fivush, R. (1993). Developmental perspectives on autobiographical recall. In: (Eds. G. Goodman & B. Bottoms), *Child Victims and Child Witnesses: Understanding and Improving Testimony*. (pp. 1-24) Guilford: New York.

Garven, S., Wood, J.M., Malpass, R & Shaw, J.S.. (1998). More than suggestion: Consequences of the interviewing techniques from the McMartin preschool case. *Journal of Applied Psychology, 83(3)*, 347-359.

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

Geiselman, R., Saywitz, K., & Bornstein, G. (1990) Effects of cognitive interviewing, practice, and interview style on children's recall performance.

Goodman, G.S., Bottoms, B.L. Schwartz-Kenney, B., & Rudy, L. (1991). Children's testimony about a stressful event: Improving children's reports. *Journal of Narrative and Life History, 1*, 69-99.

Goodman ,GS. , Batterman-Faunce, J., Kenney, R (1992). Optimizing children's testimony: Research and social policy issues concerning allegations of child sexual abuse. In: (Eds. D. Cicchetti & S. Toth) *Child Abuse, Child Development, and Social Policy*.. Norwood, NJ: Ablex, 1992,

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* (pages 249-284). NY: Cambridge University Press.

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

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

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

Hughes, M., & Grieve, R. (1980). On asking children bizarre questions. *First Language, 1*, 149-160.

Langer, E., J., & Abelson, R. (1974). A patient by any other name...Clinican group difference in labelling bias. *Journal of Consulting and Clinical Psychology, 42*, 4-9.

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

Lepore, S.J., & Sesco, B. (1994). Distorting children's reports and interpretations of events through suggestion. *Applied Psychology, 79*, 108-120.

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

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

Peterson, C., Moores, L., & White, G. (In press). Recounting the same events again and again. *Applied Cognitive Psychology.*

Pipe M.E,, Gee S,, Wilson J,G., Egerton J.M..(1999) Children's recall 1 or 2 years after an event. *Developmental Psychology. 35* 781-789.

Poole, D.A., & Lindsay, D.S. (1995). Interviewing preschoolers: Effects of nonsuggestive techniques, parental coaching and leading questions on reports of nonexperienced events. *Journal of Experimental Child Psychology., 60,* 129-154.

Poole, D.A., & Lindsay, D.S. (2000). Effects of parent's suggestions, interviewing techniques, and age on young children's event reports. *Journal of Experimental Psychology* (accepted for publication).

Poole, D., & White, L. (1991). Effects of question repetition on the eyewitness testimony of children and adults. *Developmental Psychology, 27,* 975-986.

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

Rayner, K., & Pollatsek, A. (1989). *The Psychology of Reading.* Prentice-Hall, Inc. Division of Simon & Schuster. Englewood Cliffs: NJ.

Realmuto, G., Jensen, J., & Wescoe, S. (1990). Specificity and sensitivity of sexually anatomically correct dolls in substantiating abuse: A pilot study. *Journal of the American Academy of Child and Adolescent Psychiatry, 29(5),* 743-746.

Siegal, M., Waters, L., & Dinwiddy, L. (1988). Misleading children: Causal attributions for inconsistency under repeated questioning. *Journal of Experimental Child Psychology, 45,* 438-456.

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

Walker, N., Lunning, S., & Eilts, J. (1996). Do children respond accurately to force choice questions? Paper presented to the NATO Advanced Study Institute: Recollections of Trauma: Scientific Research and Clinical Practice. Talmont Saint Hilaire, France, June.

Warren, A., & Woodall, (In press). The reliability of hearsay testimony: How well do interviewers recall their interviews with children? *Psychology, Public Policy, and Law.*

MAGGIE BRUCK AFFIDAVIT

Cause No.

EX PARTE IN THE 174TH DISTRICT COURT
KENNETH BRUCE PERKINS IN AND FOR
HARRIS COUNTY, TEXAS

I, Maggie Bruck, say as follows:

I am an Associate Professor in the Division of Child and Adolescent Psychiatry Department of Psychiatry and Behavioral Science at Johns Hopkins University School of Medicine in Baltimore, MD. I am also on leave from McGill University in Montreal, Canada where I am a Professor of psychology. I specialize in research in the field of cognitive and developmental psychology. My particular research interests focus on children's language and memory development.

I received my undergraduate degree in Psychology from Wheaton College in Norton, Massachusetts, in 1967. In 1969, I received my Master's degree in Experimental Psychology from McGill University. I earned my Ph.D. in Experimental Psychology from McGill University in 1972.

My academic/research history includes experience as a Research Associate in the McGill University Department of Psychology (1972) and at the McGill-Montreal Children's Hospital Learning Center (1972-1975). I served as a Senior Staff Member of the Center for Applied Linguistics in Arlington, Virginia, in 1975 and 1976. From 1976 through 1993 I was the Research Director at the McGill-Montreal Children's Hospital Learning Center. From 1976 through 1992 I was an Associate Member of the Department of Psychology at McGill University.

Since 1971 I have taught the following subjects: Educational Psychology (McGill); Tests and Measurement (McGill); Language Development (Sir George Williams University); Research Methods in Psycholinguistics (McGill); Childhood Psychopathology (McGill & Concordia Universities); Graduate Clinical Seminar (McGill); Psychology of Language (McGill); Graduate Cognitive Seminar (McGill); Experimental Problems (McGill); Reading Ability and Reading Disability (McGill); and Children in the Courtroom (McGill). I have lectured and organized seminars in the Medical Schools of McGill University and Johns Hopkins University..

My administrative experience includes tenure as: Research Director at the McGill University Children's Hospital Learning Center; Acting Director, McGill-Montreal Children's Hospital Learning Center; and Assistant Director of the Learning Center of Quebec, and Research Director Division of Child and Adolescent Psychiatry, Johns Hopkins University.

I have served on several McGill University committees including the Cyclical Review Committee for the Department of Otolaryngology; Graduate Faculty Social Science Research Grants Committee

(Chair), Graduate Faculty Council, Graduate Fellowship Committee (Department of Pediatrics) and Committee on Non-medical research involving human subjects.

I have also reviewed numerous grants for organizations including: Social Science and Humanities Research Council; Natural Sciences and Engineering Research Council of Canada; the McGill-Montreal Children's Hospital Research Institute; Fonds de la Recherche en Sante du Quebec; B.C. Health Care Research Foundation; National Institute of Health; and National Science Foundation.

I have reviewed articles for *Reading and Research Quarterly*; *Canadian Journal of Psychology*; *Developmental Psychology*; *Memory & Cognition*; *International Journal of Behavioral Development*; *Journal of Experimental Psychology*; *British Journal of Psychology*; *Law & Human Behavior*; *Applied Psychology* and *Applied Cognitive Psychology*.

I have served as a member of the following Review Committees: National Health and Welfare Canada Development Program; American Psychological Association (Division 7); Society for Research in Child Development; and the Advisory Committee Massachusetts General Hospital Institute of Health Professionals. I also serve on the Editorial Boards of the following scientific peer review journals: *Applied Psycholinguistics; Journal of Experimental Psychology; Psychology, Public Policy and Law; Child Development; Journal of Experimental Psychology: Applied.*

I am a member of the American Psychological Society, the Society for Research in Child Development, and the Psychonomics Society.

I have received nearly two dozen research grants during the last 20 years. I have published some 70 articles in peer reviewed publications, 20 book chapters and co-authored with Stephen Ceci, Ph.D., *Jeopardy in the Courtroom: A Scientific Analysis of Children's Testimony*, American Psychological Association: Washington, D.C. (1995). In August 2000, we will be awarded the William James Prize for excellence in psychology for *Jeopardy in the Courtroom.* Dr. Ceci and I also won the Robert Chin Memorial Award for the most outstanding article on child abuse in 1994 for our 1993 article *The Suggestibility of Child Witnesses: An Historical Review*, (published in *Psychological Bulletin, 113*, 403-439, 1993). I have also presented more than 40 peer reviewed papers at professional conferences and presented more than 50 invited addresses.

For the past decade, my research interest has focused on the accuracy of recall of past events. This work has been a collaborative enterprise with Dr. Stephen Ceci. Our research focuses on the factors that influence the accuracy of children's and adult's autobiographical memory. The results of these studies have been published in peer-reviewed journals and have been presented at scientific meetings.

I have reviewed the materials of a number of actual cases of alleged sexual abuse of children, either in the capacity of an expert witness or consultant for the defense and for the prosecution. Specifically, I have testified as an expert trial witness in *North Carolina v. Robert Fulton Kelly*; *R. v. Linda Sterling* (Saskatchewan); and at evidentiary hearings in *California v. Scott Kniffen, Brenda Kniffen, Alvin McCuan, and Debora McCuan*, and *Commonwealth of Massachusetts v Cheryl Amirault LeFave.* I have

participated as amicus curiae in *New Jersey v. Margaret Kelly Michaels* (New Jersey Supreme Court), *Snowden v. Singletary* (United States Court of Appeals for the 11th Circuit), State of New Jersey v. James Krivacska (New Jersey Superior Court), and *State v Ryan D. Smith* (Oregon, Court of Appeals). Facts and issues in these cases were similar to those present in this case.

I. EXPERT OPINION OF THE RECORD IN *PERKINS* ACCORDING TO THE RELEVANT SCIENTIFIC RESEARCH LITERATURE AND ORGANIZATION OF AFFIDAVIT

I have been asked by defense counsel for Bruce Perkins to review a number of documents (see Index) and to address the following issues:

1. Were the interviews with the child witnesses in *Perkins* so suggestive as to be capable of rendering the children's reports of abuse unreliable?

2. What factors could contribute to the unreliable reporting of the child witnesses in this case?

3. Did the experts for the prosecution provide opinions that were scientifically accurate?

The material that I reviewed was sufficient to reach the following expert opinions. **First**, the interviews in this case were indeed so suggestive as to render the children's reports of abuse unreliable. **Second,** the major factor that characterized the interviews of the children and indeed the whole investigation of the case was that of interview bias. The investigators had only one primary hypothesis—that the children were abused by Bruce Perkins. Not only did they fail to investigate or consider alternative explanations for the children's statements, but they failed to take into account any inconsistent information that would disconfirm their main hypothesis. This bias had two major results: first, it tainted the investigators own views of the facts of the case and second, in order to get the young children to provide statements consistent with these biased beliefs, the investigators used a number of suggestive interviewing techniques. **Third**, many of the crucial opinions provided by the experts for the prosecution were not based on any scientific evidence and in many cases their opinions are actually contradicted by the scientific evidence.

In this document I first focus on the concept of interview bias (Section II) because it is so central to understanding the key aspects of this case, including the investigation, the suggestive questioning of the children, and the flawed testimony of the prosecution's experts. Section II summarizes some of the relevant scientific literature one how bias influences not only the accuracy of children's reports but also the accuracy of interviewers' reports. In Section II, I also refer to the most salient details in the record and examine the degree to which interviewer bias characterized the questioning and investigation in *Perkins*.

In section III, I focus on some of the interviewing techniques used in *Perkins*. I review the scientific literature that examines the degree to which these techniques are suggestive and can taint the testimony

of young children. I then return to the record and present some examples of how the techniques were used in the interviews of the Perkins children.

In section IV, in light of the existing scientific literature, I examine the accuracy of the testimony of the prosecutions' expert witnesses. Specifically, I evaluate the scientific basis of their testimony about clinicians' ability to judge credibility of the children's statements, about the scientific literature on children's suggestibility and about their understanding of the degree to which coaching influenced the children's statements in the present case.

II. INTERVIEWER BIAS

Definition

Stephen Ceci and I have presented a model of the architecture of suggestive interviews (Ceci & Bruck, 1995). According to our model, *interviewer bias* is the defining feature of many suggestive interviews. Interviewer bias is a term used to characterize those interviewers who hold a priori beliefs about the occurrence of certain events and, as a result, mold the interview to elicit statements from the interviewee that are consistent with these prior beliefs. One of the hallmarks of "interviewer bias" is the single-minded attempt to gather only confirmatory evidence and to avoid all avenues that may produce negative or inconsistent evidence. The biased interviewer does not ask questions about the allegations that might provide alternate explanations that are inconsistent with his primary or only hypothesis (e.g., A biased interviewer would not ask, "Did your Mommy tell you that this happened or did you see it happen?"). Nor does the biased interviewer challenge the authenticity of the child's report when it is consistent with his hypothesis (e.g., The biased interviewer would not say, "It's important to tell me only what you saw, not what someone may have told you." or, "Did that really happen?", or "It's OK to say you don't remember or you don't know."). When provided with inconsistent or bizarre evidence, biased interviewers either ignore it or else interpret it within the framework of their initial hypothesis. It is important to note that within this context, a biased interviewer may be a police officer, a therapist, and even a parent. It takes no special skills to be a biased interviewer.

Scientific Studies of Interviewer Bias

Interviewer bias has been the focus of much study in the area of psychology. The general findings of this research are that biased interviewers can elicit a number of false reports from the person being interview. Also biased interviewers will inaccurately report the contents of the interview; they will sometimes recall their own a priori beliefs rather than the statements of the interviewee. The following three studies give a flavor of the research. The first study focuses on bias and clinical judgments; the second and third studies focus on the effects of interviewer bias on young children of the same age as those in *Perkins*.

Clinical judgments and *interviewer bias.* Langer and Abelson (1974) showed a videotaped interview of a man to two groups of therapists with different theoretical orientations. One group included behavior therapists who were trained to focus on the patient's presenting symptoms and to find methods to modify or alleviate these. The second group included therapists with a traditional psycho-dynamic orientation which calls for the understanding and interpretation of symptoms in term of childhood experiences or inner conflicts. Although all therapists were to view the same videotape, half were told that the man was applying for a new job and the other half were told that the man was a patient undergoing a psychiatric interview. After viewing the videotape, all the therapists were asked for judgments of the man's adjustment and behaviors. Psycho-dynamic therapists (who have a priori beliefs about the origins and meanings of behaviors) gave more positive ratings when told that man was a job applicant rather than a patient, even though exactly the same videotape was shown. For example, these therapists characterized the job candidate as: attractive, candid, fairly open, whereas they characterized the patient as tight, defensive, passive-aggressive, having a conflict over homosexuality. This trend was not shown by the behavioral therapists who were trained to only pay attention to the behavior without any attempt of interpretation (i.e., a priori beliefs). The behavioral therapists characterized the job candidate and the patient in the same manner.

Interviewer bias influences *children's reports.* Thompson, Clarke-Stewart & Lepore (1997) conducted a study in which 5- and 6-year-olds viewed a staged event that could be construed as either abusive or innocent. Some children interacted with a confederate named "Chester" as he cleaned some dolls and other toys in a playroom. Other children interacted with Chester as he handled the dolls roughly and in a mildly abusive manner. The children were then questioned about this event. The interviewer was either 1) "accusatory" (suggesting that the janitor had been inappropriately playing with the toys instead of working), 2) "exculpatory" (suggesting that the janitor was just cleaning the toys and not playing), or 3) "neutral" and non-suggestive . In the first two types of interviews, the questions changed from mildly to strongly suggestive as the interview progressed. Following the first interview, all children were asked to tell in their own words what they had witnessed and then they were asked questions about the event. Immediately after the interview and two weeks later, the children were asked by their parents to recount what the janitor had done.

When questioned by a neutral interviewer, or by an interviewer whose interpretation was consistent with the activity viewed by the child, children's accounts were both factually correct and consistent with the janitor's script. However, when the interviewer was biased in a direction that contradicted the activity viewed by the child, those children's stories quickly conformed to the suggestions or beliefs of the interviewer. Also children's answers to interpretive questions (e.g., "Was he doing his job or just being bad?") were in agreement with the interviewer's point of view, as opposed to what actually

happened. When asked neutral questions by their parents, the children's answers remained consistent with the interviewers' biases.

The development and consequences of interviewer bias. Bruck, Ceci , Melnyk, & Finkelberg (1999) showed how interviewer bias can quickly develop in natural interviewing situations, and how it not only taints the responses of child interviewees but also the reports of the adult interviewers.

In this study, a special event was staged for 90 preschool children in their school. In groups of three and with the guidance of research assistant A, the children surprised research assistant B, played games, ate food, and watched magic tricks. Another 30 children did not attend the birthday party but in groups of two, they simply colored a picture with research assistants A and B. These children were told that it was one of the assistant's birthday.

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

Bruck et al. found that the fourth children (those that attended the coloring event and who were interviewed last) produced twice as many errors as the children who attended the birthday party; 60% of the children who only colored, made false claims that involved a birthday party. This result suggests that the interviewers had built up a bias that all the children had attended a birthday party. By the time they interviewed child # 4, the interview was structured in such a way as to elicit claims consistent with their primary hypothesis. Thus if interviewers have the belief that all the children they are interviewing have experienced a certain event, then it is probable that many of the children will come to make such claims even though they were non-participants (or non-victims). Another important finding was that even when child #4 denied attending a birthday party, 84 % of their interviewers later reported that all the children they interviewed had attended a birthday party. These data suggest that regardless of what children actually say, biased interviewers inaccurately report the child's claims, making them consistent with their own hypotheses.

These studies and others reviewed by Ceci & Bruck (1995) provide important evidence that interviewers' beliefs about an event can influence the accuracy of children's answers. The data highlight the dangers of having only one hypothesis about an event, particularly an event involving an ambiguous act such as touching.

Interviewer Bias in Perkins

In the context of *Perkins*, examples of interviewer bias appear at all stages of the investigation, including the initial questioning of the children by their mothers, the later questioning of the children by police and therapists, and the manner in which the investigation was carried out. During all aspects of the case, it appears that the major hypothesis was that Bruce Perkins had abused his grandchildren. Even when glaring disconfirmatory evidence was presented, it was disregarded. Although there is always an element of interviewer bias in most investigations, the degree of bias in the present one was extreme. It was so strong that it could account for the emergence of allegations of sexual abuse by the young child witnesses. The following examples are a few among the many instances of how interviewer bias was instantiated in *Perkins*.

Bias of the parents. By the late summer of 1992, Trish Perkins developed a suspicion that Bruce had sexually abused her children, Josh and Jaclyn. Based on these suspicions, she began questioning her 4-year old daughter Jaclyn who eventually made allegations that were consistent with her mother's beliefs. As a result of these allegations, the other two children, Mandy and Josh, were specifically questioned about abuse.

There were three incidents that aroused Trish's suspicion that her children had been abused. First, in late August 1992, Trish observed Bruce and Josh sitting on a bed watching a videotape. Trish did not like the way Bruce was sitting with Josh (Tr. 701); nonetheless, she did not try to separate them or express her disapproval. The second incident which occurred near the same time, involved Bruce taking Josh on a tractor ride. Trish did not halt the ride and she did not know how Bruce was holding Josh, but nonetheless, she felt that this portended wrong doing (Tr. 723). Finally, Trish became "suspicious" when Bruce and Carol sent their grandchildren a card with "going-back-to-school" money". Although the grandparents had consistently given the children money over a period of time, Trish told Detective Bynum that she was obsessed that "these people are paying my children to be silent. I can't get away from it, I feel like I am crazy but I can't get away from it". (Transcript Bynum interview with Trish Perkins, p.2). Not only did Trish obsessively think that Bruce had abused her children, but she directly asked her four-year-old daughter about this belief: "Why is Papa Bruce sending you money?" (Tr.710) and "Has Papa Bruce touched you in a way that you don't like, has he touched your private parts? " (Bynum p. 2). Although Jaclyn denied abuse, Trish continued the questions until Jaclyn made allegations consistent with Trish's belief. Trish told her brother-and-sister-in-law about Jaclyn's allegations and the biased questioning began with their four-year-old child, Mandy.

This was not, however, the first time that Trish Perkins suspected that her children had been abused. A review of the record reveals that her bias in 1992 had a long history. As early as September 1991, one year before the first allegation and one month before the alleged abuse at Bruce's birthday party (October 6 1991), Trish asked her three year old daughter Jaclyn if a neighborhood child had

touched her in a "mean way." (Kammholz, 402) She asked her at least twice on the same day (Tr. 1696).1 Jaclyn assented and made other claims of sexual play with the neighborhood boy. However, it does not appear that there was any explicit investigation of the Jaclyn's claims.

Between September 1991 and September 1992, Trish suspected a number of candidates who were abusing her children. These included: Jennifer Pierce (Tr. 538), Mandy Perkins (Tr 1700), and even her son Josh (9/10/92 Bynum interview p. 9). At various times Trish also thought that Jennifer Pierce had been abused (Tr.734), that Sean Rothrock had been abused (Tr 1695)) and that Mandy Perkins may have been abused (Tr. 1700). These suspicions were based on her seeing her children play "doctor" with other children. Although she never observed any explicit sexual activity, she did see her children playing naked and inspecting their playmates bodily parts (Tr. 412). In fact her suspicions and concerns were so great that all the Perkins children were evaluated by a psychologist, Kammholz.

Trish's suspicions and subsequent accusations may also reflect aspects of her own history where indeed there were three major instances of allegations and accusations made against family members. First, Trish's younger sister Diane accused Larin of making a pass at her (Berliner notes September, 1991). Second, Trish was reportedly raped when she was 17 years old (Berliner notes September, 1991). Third, Trish was concerned that she may have been abused by her own grandfather (Tr 1079, 1084). These facts raise the hypothesis that Trish's belief that her children were abused may have initially stemmed from Trish's own experiences rather than from any behaviors of her own her children.

Biased Investigator Bynum:. The formal investigation began on September 8, 1992. Patty and Lann Perkins along with their 4-year old daughter, Mandy, were interviewed by Detective Bynum. The next day, Bynum interviewed Larin, Trish, and Jaclyn. These initial interviews provide many examples of interviewer bias that eventuated in the indictment of Bruce Perkins for sexual abuse. The following examples taken from these first interviews provide highlights of the bias.

First, Bynum obtained the details of the alleged abuse from the parents and then used this information to structure his interviews with their children. That is, the children were never asked to tell in their own words if anything had happened with their grandfather. Rather the interviews with the children were structured to get the child to confirm the details provided by their parents. The structure of the interviews did not allow the children to provide evidence that might contradict or provide alternative hypotheses for their parents' reports.

---

1 There is some confusion in the record about the timing of this accusation and questioning. Paula Porter testified, "She came to my house last spring of 1992 and thought that Jaclyn had been molested. She had just found this out the night before that morning" (Tr. 1695). Trish told Porter that she had asked Jaclyn if anyone had been mean to her. But the dating is in error because around the same time (March, 1992) Trish told her child's therapist, (Kammholz) that she had questioned Jackie about the neighborhood child in September of 1991 (Tr. 402). So although Trish told her neighbor that she had just found out about molestation in the spring of 1992, she had told a therapist that she had found out almost 6 months previously. This discrepancy raises concern about the motivation of Trish's reporting error and the reliability of her testimony.

Second, Bynum used certain procedures that are highly contaminating and would not be used by any skilled interviewer. The children were in earshot when Bynum interviewed their parents and obtained details of the allegations (Tr. 1257). Both mothers were allowed to be present when Bynum interviewed their children. The mothers were allowed to prompt their children, reward their children, and provide their children with answers. At trial, Bynum did not consider the possibility that these two procedures could taint the evidence (Tr. 1202).

A third marker of bias in Bynum's investigation was that he built a case against Bruce despite the fact that the children had provided few coherent details or knowledge of sexual abuse.

EXAMPLE

Mandy Perkins did not make any spontaneous statements about sexual abuse in the first interview. She only supplied (often one word) responses to very specific questions. Although she reported that Bruce touched her vagina she denied the essential details of sexual abuse: seeing or touching the alleged perpetrator's penis. It is difficult to imagine how sexual abuse could take place without these crucial elements. Bynum overlooked this fact and did not question the child further.

EXAMPLE

When Jaclyn was asked in different parts of the interview to provide names of those who had molested her, she named everyone but Bruce.

BYNUM: Has anybody ever done anything to you that was not right
JACLYN: Yes
BYNUM: A big person
JACLYN Yes
BYNUM Who?
JACLYN: Jennifer (Bynum/Jaclyn transcript of tape, p. 10)

Finally with more prompting, Jaclyn named Bruce. Obtaining the desired answer, Bynum proceeded to a new line of questioning.

Fourth, it is clear from the transcripts that the children had difficulty understanding the Bynum's questions (which were sometimes poorly formulated for 4 year old children). At times, the children had difficulty expressing themselves and producing a coherent answer. An unbiased and competent interviewer would establish that the child understood the questions. An unbiased interviewer would ask for clarification to ensure that the interviewer's interpretation of the children's often ambiguous statements were consistent with what the child meant to say. Bynum did none of the above. Rather he inserted his own interpretations which were consistent with the theme that Bruce Perkins abused his grandchildren.

EXAMPLE:

BYNUM: But you said Papa Bruce has touched your vagina before, has he done it more than one time? You are shaking your head yes, has he done it?
MANDY: But Jennifer did, Paw-Paw Bruce and Jacob.

BYNUM : Jennifer, Papa Bruce and Jacob touched you there too, on the vagina. (Bynum/ Mandy tape transcript p. 6)

Comment: It seems that Mandy had difficulty constructing a response and produced disconnected fragments of a sentence that Bynum directly interpreted (without asking, "Is this what you mean"?) as Bruce and the other children touched Mandy.

Example

BYNUM: Well did he ever put anything else on your vagina?

MANDY: Un-un
BYNUM: Uhhhhhh, did he ever put anything wet?
MANDY Uhhh, no
BYNUM: How about, did he ever use any kind of oil or ketchup?
MANDY: Ketchup
BYNUM: He has put ketchup on your vagina. (Bynum/ Mandy tape transcript p. 10)

Comment: Mandy answered inconsistently—first she said there was nothing wet and then there was ketchup. This discrepancy was not pursued. Furthermore the third question about ketchup and oil is not only bizarre (although it does reflect the investigator structuring the interview to confirm the parents' hearsay reports) but it is incomplete; the full intent of the question is not explicitly stated. Bynum meant to ask "Did he ever use ketchup on your vagina", but he omitted the crucial part of the question "on your vagina". Four-year old children will not necessarily know that "on the vagina is implied" . Thus Mandy may have thought that she was being asked if Bruce ever used ketchup. However, given the biases of the interviewer, it was interpreted as another detail of her abuse confirming her parents' reports.

Fifth, Bynum overlooked the children's inconsistent statements which would have been red flags to unbiased interviewers who would have questioned the children further about these statements. Bynum just let them pass.

EXAMPLE:
MANDY: Jaclyn needs some help.
.....
BYNUM; ...Jaclyn needs some help, why, cause somebody has done something to her too?
MANDY: Papa Bruce has
BYNUM: Have you ever seen him do that to Jaclyn?
MANDY: No
BYNUM: .....How do you know that?

(COMMENT: This is one of the few examples in the case where the interviewer seemingly used a neutral unbiased technique)

MANDY: Cause my momma told me.
BYNUM: ...OK, when Papa Bruce touched you ...(Bynum/ Mandy tape transcript p. 10)

COMMENT: Mandy's statement that she never saw Jaclyn being abused but was only told by her mother caused no obvious concern for Bynum who continued to question the child about being molested by Bruce.

Finally, Bynum's biases were reflected not only in the suggestive interviews with Mandy and Jaclyn, but also in terms of his conduct through out the investigation. For example, Bynum appeared to let Bruce Perkins' sons and daughters-in-law dictate some of the crucial aspects of the investigation. They encouraged Bynum not to interview Josh (the oldest of the children) although they did believe that he had been abused (Tr. 1184). It is also concerning that there were no CPS interviews, a standard procedure in investigations of child abuse. Perhaps Bynum was persuaded to disregard the normal protocol because he was so persuaded that the children were abused. As a result the only official investigatory interviews are those of Jackie and Mandy with Bynum.

In addition, the majority of Mandy's and Jaclyn's claims that other children had witnessed the abusive episodes were not substantiated. Jennifer Pierce, Jacob Pierce, and Katie Pruitt all denied being present. (Tr 1191, 1194, 1196) Travis Smith denied any knowledge or participation(Tr. 1197)--an accurate statement since he did not move into the neighborhood until some time later. Even though, Mandy's and Jaclyn's descriptions of central events (i.e., who was there, what other people did) were inaccurate the prosecution nonetheless indicted Bruce Perkins for the sexual abuse of his grandchildren.

Bias of the therapists. Three therapists testified for the prosecution that Mandy, Jaclyn and Josh had all been abused by their grandfather. Their bias is reflected in a number of ways. The therapists uncritically accepted the parents' conclusions that Bruce Perkins had abused their children. The therapists interpreted all of the children's behaviors in and out of the therapy situation as symptoms and consequences of childhood sexual abuse, The therapists did not entertain plausible alternatives for the children's behaviors, such as the birth of a sibling, exposure to material with sexual themes, the children's own fear of their being accused of sexual abuse. Finally, the therapists supported their opinions with subjective clinical observations, with pseudo-scientific evidence, and sometimes with inaccurate summaries of the scientific evidence. In this sense, the therapists went beyond the pale to support their biased opinions.

The therapists' defended their biased practices by stating that they were not evaluators but therapists (Tr. 135, 842, 1152). Although this is a position taken by many therapists in cases similar to *Perkins*, the statements reveal poor clinical practice and reflect the degree to which bias operates in these cases. In *Perkins*, this stance allowed the therapists to simply accept the parents' statements that the children were abused without any further probing of the situation. By simply relying on the "truth" of the parents' statements, the therapists could interpret all the children's behavior as consistent with sexual abuse, no matter how bizarre the behavior. There are major problems however with this practice. First, it assumes that the parents' reports were completely accurate. In *Perkins*, the therapists seemed bound to the word of the parents, never questioning their accounts, and never attempting to gather more background information in order to better understand the children' problems. Second, the notion that a therapist does not have to conduct an independent evaluation can have devastating consequences as may

have happened with the young Perkins children. Namely, the therapists treated the children for a presumed trauma. They encouraged the children to talk about the trauma, to show the trauma in their play, and to confront the trauma. Over the period of treatment, at least two of the children Jaclyn and Josh, made no progress or regressed. One major hypothesis for this decline is that the children were receiving treatment for the wrong "illness". More primary causes of their symptoms were ignored but still remained in play thus prolonging their behavioral problems. Competent therapists would continuously collect information and test alternative hypotheses when children's problems remained static or worsened.

A few of the many examples of bias for each of the therapists are now described.

Margaret Richardson-McCollum (MRM) was Jaclyn Perkins' therapist. MRM's one and only guiding hypothesis in treatment was that Jaclyn was sexually abused. Signs that the child's report might be unreliable that might cause unbiased therapist to consider alternative hypotheses, caused no concern for MRM who interpreted this discrepant evidence as yet another sign of sexual abuse.

EXAMPLE.

Jaclyn told MRM: "My brother came in and he saw and he told my mom to come in and saw." (Tr. 986). When MRM found out that Mom was not present, MRM explained that Jaclyn's statement reflected her anger at her mother (she interpreted the child's statement of one of anger against her mother (Tr. 1008).

Comment: . This is a rather convoluted conclusion. A more straightforward hypothesis might be that Jaclyn was not a reliable reporter.

EXAMPLE.

Jaclyn's behavior did not improve but rather deteriorated over the course of therapy. She began talking of bad devils and having nightmares (Tr. 1022, 1031). Even though MRM testified that the deterioration was due to Jaclyn's fear of going to court, she admitted that she had never asked Jaclyn about the root of her fear (Tr. 1030)

COMMENT: When patients are treated for problems that they do not have it is often the case that their symptoms will remain the same or else deteriorate. In the case of Jaclyn, it is possible that she was being traumatized by having to talk about crimes that never happened. Her fear resulted for the implanted belief that she had been abused. MRM did not consider this explanation.

EXAMPLE.

MRM interpreted all of Jaclyn's behaviors as consistent with sexual abuse even when she said nothing happened or that she did not remember. MRM testified that these patterns reflected Jaclyn's difficulty producing details of sexual abuse , that when the topic became too emotional she would withdraw into silence and that she needed much guidance to produce disclosures (Tr.992, 1031, 1146.

Comment:. These behaviors could be a sign that the child had nothing to say but MRM's only interpretation was that her discomfort and denial was a sign that she was getting too close and thus frightened of disclosing her abuse.

In general MRM and the other therapists could not scientifically justify their treatment techniques and the basis of their clinical judgments. But because the defense never asked for the basis of their opinions and never challenged their answers, the jury did not understand the illogical and biased nature of their conclusions.

Connie Nelke (CN) was Mandy Perkins' therapist. CN never questioned the validity of the mother's statement that Mandy had been abused and she never questioned the accuracy of the details of abuse provided by the mother. This position demoted the status of the therapist who became an agent of the mother. Her job became one of interpreting all behaviors in terms of sexual abuse.

EXAMPLE

CN's belief that the children were sexually abused was based (in part) on the mother's report that the children engaged obsessively in sexually explicit play. Although Nelke had never observed sexualized behaviors, she testified that Mandy's sexual acting out was consistent with sexual abuse (Tr. 823). However, only at trial does Nelke learn that the mother also never saw the child engaged in explicit sexual play Tr. 837).

Comment: : CN's failure to conduct an adequate intake of the parents' reports, her failure to explicitly question the parents about the basis of their claims (other than my child told me) led to an inadequate and possibly incorrect conceptualization of the problem.

The following example shows how Nelke passively allowed the mother to control the therapy sessions and the interpretations of what occurred in the sessions.

CN testified, "Mandy drew a picture of Josh, Jaclyn, and sour cream. She (Mandy) said he put sour cream on our bottoms. Here's the sour cream in a little container... That was confusing to me,.... and she hadn't talked about this before."(Tr. 805)

Comment: Instead of trying to resolve her confusion by asking Mandy to elaborate about the meaning of her picture, Nelke allowed the confusion to stand and to be resolved by Mandy's mother

CN's testimony continued "Her mother reported to me that Mandy had cried after the last session because she had lied to me about the sour cream, that the sour cream wasn't in a container in a refrigerator but it was from his penis but she was embarrassed to tell me because she didn't like to talk about it. The mother said that Mandy reminded the mom prior to this session to tell me she had not told the whole truth." (Tr 807). Mandy then told CN that "The sour cream came

from his penis. He told me it was magic. He would put it on us with his hands, on our bottoms. I think catsup came from his penis too." (Tr. 807).
Comment: An unbiased clinician would not uncritically accept the mother's explanation. The first obvious question would be how and why did Mandy's mother find out in such detail what had happened in the therapy session. An unbiased and skilled child therapist would also question the capability of a four-year old child to spontaneously express remorse. The unbiased clinician would ask if the mother had attempted to taint the child's statements.

Robert McLaughlin (RM) was Joshua Perkins' therapist. RM's major task was to explain to the jury (a) why Joshua denied abuse for such a long period of time; after 10 months of therapy, Joshua was only beginning to disclose details of abuse to RM; (b) why Joshua had never disclosed any threats in the therapy room and (c) why Joshua appeared normal outside of the therapy room but was tortured inside the room. Joshua did not want to go to therapy, often he was carried in by one parent, often he cried until he was carried out. (Tr. 142).

In general, RM testified that Joshua's denials, silence, and avoidance of the therapy room were signs of abuse. Not knowing the answers to questions reflected a reluctance to admit abuse (Tr. 126, 127, 143) Simply put, when Joshua said "No" it meant "Yes".

RM invested in this position because like the other therapists, he accepted the referring statements that Joshua's grandfather had abused him and threatened him not to tell of his abuse. RM accepted this as historical truth and as a result made the following illogical or circular diagnoses.

RM testified that Joshua was suffering from Post Traumatic Stress Disorder. "It includes a number of symptoms. A key one with regard to my impressions of Joshua has to do with avoidance of particular stimuli that appear to be related to whatever the traumatic event might be. In this particular instance, since what I have been told have (sic) to do with abuses from his grandfather, what he is avoidant of in the therapy session on many occasions are references to his grandfather where he seems to become very depressed and inhibited, very thoughtful and withdrawn at those times" (TR. 124). In other words, RM has been told that Joshua was abused by his grandfather. Therefore his avoidance of the topic must reflect his reaction to the trauma.

Both RM and MRM had to explain to the court why Joshua excelled socially and academically in school and yet was so dysfunctional in the therapy room. Both provided the same circular explanation: "Whatever their typical coping mechanisms, (they) get stronger, but in areas where they are not able to cope, they decompensate. So the child who has a lot of academic skills, will (put), a lot of energy into that... they may continue to perform well academically if that's their area of strength. (TR. 151). Of course, RM's opinion does not explain Joshua's extreme avoidant behaviors in the therapy room and again his explanation is based on the central premise that Joshua was abused.

The following list of alternative hypotheses were not entertained by Joshua's therapist

-Joshua doesn't talk about abuse because there is nothing to talk about.
-Joshua's problems in the therapy room reflect parental (and maybe therapist's) pressures to disclose false information.
-Joshua is concerned that, like his grandfather who went to jail, he may be accused of sexually abusing his sister (a concern that his mother related to the police in the early part of the investigation)
-Joshua is at peace in school because this is the one place where he is not being pressured to falsely disclose

The fact that RM presents no details of disclosures at trial, the fact that RM only knows about threats from Joshua's parents but has never been told by Joshua and the fact that Joshua overheard his parents reminding the therapist of the nature of their concerns provides some support for these alternative hypotheses.

However these alternative hypotheses are rejected by RM who assured the court that because of his superior clinical skills there could be no errors in diagnosis (Tr. 144). As discussed below, RM also misinterpreted the scientific literature to support his unscientific and highly biased opinion.

Summary

The belief that Bruce Perkins had abused his grandchildren was initially generated by the parents of the child witnesses. The parents for whatever reasons seemed to foster and transmit this belief to the police (who carried out an incomplete investigation) and to the therapists (who carried out no investigation but just accepted the belief). Although the children said little throughout the case to substantiate these beliefs, all that they did say and do was interpreted as evidence to confirm abuse.

With time, the children's narratives and allegations became more consistent with sexual abuse. It is possible that they needed the time, the comfort, and perhaps even the cognitive resources to tell their horrible stories. However, it is also as plausible that their stories were simply false and the reflection of the biases of their investigators, parents and therapists. In the next section, I discuss how biases become translated into suggestive techniques that can taint the accuracy of children's reports.

SUGGESTIVE INTERVIEWING TECHNIQUES

In the past decade, numerous studies have examined the interviewing strategies that can lead young children to make false reports about central events, especially those that involve their participation and that involve bodily touching or insinuations about sexual abuse. Some studies have examined the influence of certain interviewing techniques that were used with the Perkins children, such as the use of anatomically detailed dolls and the repeated use of questions within and across interviews. As will be described in this section, the use of these suggestive techniques, especially in the hands of biased

interviewers, can bring young children to make claims about events that they have never experienced. In some studies, these false claims involve potentially serious actions that sometimes involve false claims of inappropriate actions to their bodies and that sometimes have a sexual interpretation.

The first part of this Section contains a summary of the scientific data on some of the many suggestive strategies that were used in interviews with the Perkins children. The second part of this section provides examples of how these strategies were used in the interviews with Jaclyn, Mandy, and Joshua. In providing examples, I am limited by the fact that the only electronically recorded interviews (that provide verbatim information) were Bynum's initial interviews with Mandy and Jaclyn Although the therapists testified from notes of their therapy sessions, these are often inaccurate because they omit information, or because the information is incorrectly recorded. Of most importance for the purposes of this report, the written reports do not contain every question asked by the interviewer or every answer given by the child. This information is crucial in order to evaluate the degree to which different suggestive techniques were used in each interview. Even with this limitation, there are sometimes substantial hints from the testimony that provide a flavor of how the children were interviewed, how the children reacted and what the children said.

Scientific Studies of Suggestive Interviewing Techniques

1. *Open-ended vs. Specific questions.* In order to obtain confirmation of their suspicions, biased interviewers may not ask children "open ended" questions such as, "What happened?", but instead quickly resort to a barrage of very specific questions, which require the child to provide one word answers (e.g.,. yes or no). Sometimes the questions are very leading (for example, asking the child "Where did Bruce touch you" is very leading if the child never mentioned touching by Bruce) and sometimes the questions are repeated until the child provides the desired response.

Although the strategy of asking specific questions, leading questions, and of repeating questions ensures that the child will provide information, it is also problematic because children's answers to these types of questions are often inaccurate. For example, Peterson and Bell (1996) interviewed children (ages 2 to 5 years) after they had been treated in an emergency room for a traumatic injury. They were first asked free recall questions ("Tell me what happened"). Then in order to obtain additional information, the children were asked more specific questions (e.g., "Where did you hurt yourself?" or "Did you hurt your knee?"). Peterson and Bell found that children were most likely to accurately provide important details in free recall. Across all age groups, errors increased when children were asked more specific questions; the percentage of errors elicited by free recall and specific questions was 9%, and 45% respectively.

Forced choice questions (e.g., "Was it the man or the woman?") also compromise the reliability of children's reports. This is because children will select a response set (for example, some will frequently select the second rather than the first option) and because children commonly do not provide "I don't

know" responses (e.g., see Walker, Lunning, & Eilts, 1996) even when the question is nonsensical (Hughes & Grieve, 1980). One of the reasons that children so willingly provide answers to specific yes/no or to forced choice questions even though they may not know the answer is that young children are cooperative: they perceive their adult interviewer as truthful, and not deceptive. In order to comply with a respected adult, children sometimes attempt to make their answers consistent with what they see as the intent of the questioner rather than consistent with their knowledge of the event (see Ceci & Bruck (1993) for a review). Because of this compliant cooperative characteristic, and because of young children's poor performance on specific questions, it is particularly important in interviews to tell children that they have the option of saying "I don't know" or "I don't remember".

As detailed below, when first interviewed by Detective Bynum, the Perkins children were not asked open-ended questions, rather they were asked dozens of leading questions. The children were never told that it was alright to not know the answer to the questions.

• 2. *Repeating Interviews and Repeating Questions.* In formal investigations, children are often interviewed on many different occasions. Although there are numerous concerns about the influence of these repeated interviews on children's reports, there is the most concern when these repeated interviews are conducted by biased interviewers. In these situations, children are repeatedly exposed to leading (and perhaps) misleading questions. The results of several studies indicate that when this occurs, children's reports become highly tainted, sometimes after just a few interviews and sometimes after just a few questions.

For example, Bruck, Ceci & Hembrooke (1997), repeatedly interviewed preschool children about two true events and two false events. The two true events involved the child helping a visitor in the school who had tripped and hurt her ankle and a recent incident where the child was punished by the teacher or the parent. The two false events involved helping a lady find her monkey and witnessing a man steal food from the daycare. In the first interview children were simply asked if each event had ever happened to them. If they said yes they were asked to describe the event. During the next three interviews, the children were suggestively interviewed (e.g., they were asked repeated leading questions, they were praised for responses, they were asked to try to think about what might have happened; they were told that their friends had already told and now it was their turn). In the fifth interview, a new interviewer questioned each child about each event in a nonsuggestive manner. Across the five interviews, all children consistently assented to the true event about helping a lady who fell in the daycare. However, children were at first reluctant to talk about the punishment event; many of the children denied that the punishment had occurred. With repeated suggestive interviews, the children agreed that the punishment had occurred. Similar patterns of disclosure occurred for the false events; that is, children initially denied the false events but with repeated suggestive interviews they began to assent to these events. By the third interview, most children had assented to all true and false events. This pattern continued to the end of the experiment.

One of the rationales for re-interviewing children is that it provides them additional opportunities to report important information that was forgotten or simply not reported in earlier interviews. Thus, it is assumed that when children provide new details in subsequent interviews, that these new reports are accurate memories that were not remembered in previous interviews. However, the results of recent studies dispute this claim. One set of studies consistently shows that reports that emerge in a child's first interview with a neutral interviewer are the most accurate. When children are later interviewed about the same event and report new details not mentioned in a previous interview, these have a high probability of being inaccurate whether children are interviewed by a neutral interviewer (Peterson, Moores & White, in press; Pipe et al., 1999;) or by a very suggestive interviewer (Bruck et al., 1997). Thus insertion of new but inaccurate details can be a natural memory phenomenon, it can be due to prior suggestions that become incorporated into later reports, but it can also be due to the demand characteristics of the interview. When interviewers urge children to tell them anything (that is consistent with abuse), these requests for additional information will often result in false reports which are supplied to comply with the interviewer's wishes to tell more (e.g., Bruck et al., 1997). The probability of obtaining false reports also increases as a function of the delay between an alleged event and a later interview (e.g., Bruck et al., 1995a). This is because children's memory of the original event (e.g., remembering what happened at Papa Bruce's birthday party over a year ago) fades after a period of time, allowing the suggestions (Papa Bruce did bad things) to become more easily implanted.

The studies just reviewed show the detrimental effects of repeated questions across interviews. However, the same results are obtained in the first interview when children are repeatedly asked suggestive questions. Biased interviewers sometimes repeatedly ask the same question until the child provides a response that is consistent with their hypothesis. Poole and White (1991) found that that asking the same question within an interview, especially a yes/no question, often results in young children changing their original answer.

In another study, Garven, Wood, Malpass, & Shaw (1998) found that preschoolers provided increasingly inaccurate responses to misleading statements and questions as a suggestive interview proceeded. In this study, children were suggestively interviewed for 5 to 10 minutes about a stranger who came to read their class a story. As a result of the suggestive devices, children falsely claimed that the visitor said a bad word, that he threw a crayon, that he broke a toy, that he stole a pen, that he tore a book and that he bumped the teacher. Importantly, the children came to make more false claims as the interview progressed: that is within a short 5 to 10 minute interview, children made more false claims in the second half than in the first half of the interview.

The finding that children change their answers to repeated questions is thought to reflect their social compliance and sensitivity to the demands of the interviewer. Thus children change their answers because they reason that the question was asked a second time either because the first answer was wrong or because the interviewer must not have like the first answer, regardless of its accuracy (e.g.,

Siegal, Waters, & Dinwiddy, 1988). Thus children cooperate by changing their answers and they quickly learn the types of answers that their interviewers want to hear.

Many of the patterns of behavior shown by the Perkins children match those of children in laboratory studies. The Perkins children were interviewed repeatedly about sexual abuse. The questioning began as early as October 1991. The initial interviews were conducted by their parents and only after many sessions of questioning did the children provide allegations. Also, within some interviews, the children's allegations increased as the interview continued. Although it could be argued that this reflects the child's growing comfort and willingness to open up, it can also be argued that the children's allegations increased as a result of repeated questions and as a result of their sensitivity to answer the questions in a certain way.

• 3. *Anatomically detailed dolls* are frequently used by professionals when interviewing young children about suspected sexual abuse. The major rationale for the use of anatomical dolls is that they allow children to manipulate objects reminiscent of a critical event, thereby cueing recall and overcoming language and memory problems as well as motivational problems of embarrassment and shyness. However, research conducted in the past decade has raised concerns about the suggestiveness of the dolls and their influence on the accuracy of children's reports. There are several important findings of this research. First, there is no consistent evidence to suggest that there are characteristic patterns of doll play for "abused" children. Many studies show that 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. Second, more recent studies indicate that use of the dolls does not improve accuracy of young children's reports and in some cases they decrease accuracy. For instance, we found that three year old children (Bruck, Ceci, Francoeur & Renick, 1995) and four year old children (Bruck, Ceci & Francoeur, 2000) who had just completed a medical examination at their pediatrician's office made a number of errors when asked direct questions about where the pediatrician had touched them and that these errors increased when children were asked these same questions in conjunction with dolls. Specifically, children inaccurately showed that the doctor had touched their genitalia or buttocks when this did not happen. These inaccurate answers reflect the novelty of the dolls which prompted the children to explore the genitalia, often in very creative ways; the inaccurate answers also reflected the implicit demands of the interview which were to show and talk about touching.

There is also a single case study that suggests that repeated exposure to the dolls may lead young children to fabricate highly elaborate accounts of sexual abuse. After a third exposure in a period of a week to an anatomically correct doll, a nonabused 3-year child told her father that her pediatrician had strangled her with a rope, inserted a stick into her vagina and hammered an earscope into her anus (see Bruck et al, 1995).

Thus the risk of using dolls with preschool children is well supported by the published scientific evidence--so much so that the following statements have been made by independent groups of researchers and experts about the interviewing of young children:

> "Particular caution is called for when interpreting the reports of children aged 48 months and younger, who are asked repeated leading and misleading questions about touching in a doll-centered interview" (p.218) [Koocher, Goodman, White, Friedrich, Sivan & Reynolds (1995)
>
> "It may not be appropriate to use AD dolls with either preschoolers or children who have the cognitive capacities of typical 2 to 4 year olds" (p. 190). And "Interviewers need to realize that the benefits they expect from using dolls can easily be outweighed by questions regarding the reliability of the children's reports" (p. 195). (Poole & Lamb, 1998).

The Perkins children were repeatedly interviewed with anatomically detailed dolls. Kammholz used dolls in her therapy sessions, Bynum used dolls in his investigative interviews, and the therapists sometimes used dolls. Each of these occasions could promote the elaboration of false allegations of touching, as suggested by the scientific literature.

- 4. *Other Suggestive Techniques* were used in Perkins such as the use of peer pressure (telling the child that another child had already told about abuse),. selective reinforcement (rewarding the child verbally or nonverbally for appropriate responses and ignoring responses that do not focus on abuse) and emotional atmospherics (telling the child that they are brave, not to be afraid to tell). Each of these has been found to increase the errors in children's reports when children are interviewed by biased adults (see Ceci & Bruck, 1995). As will be shown, these techniques were used to interview Joshua, Mandy and Jaclyn Perkins.

By themselves each of the techniques if used one time might have little influence on the accuracy of children's reports. But when these are used repeatedly in combination with other suggestive interviewing techniques, they can have devastating effects on the accuracy of young children's reports (e.g., Bruck et al., 1997; Garven et al., 1998, in press).

Summary of Major Scientific Findings on Children's Suggestibility

1. It is crucial to examine the conditions under which children make their first disclosures of an event. On the one extreme, if children spontaneously make claims of abuse without any previous questioning, evidence of bias or motive to falsify information, then one can have the highest faith in the accuracy of such reports. However, if children first deny the occurrence of an event and their disclosures evolve as a result of one or more suggestive interviews conducted by biased

interviewers then there is a high risk that their reports are inaccurate and tainted by the interviewing situation. The latter situation best describes the conditions in *Perkins*.

2. Suggestive interview techniques are not limited to the use of misleading questions. Rather, a range of interviewing techniques can negatively influence the accuracy of children's reports. These include verbal dimensions of the interview (e.g., the way questions are asked, the number of times questions are repeated, the use of reinforcement/punishment, etc.) but they also include nonverbal techniques such as the use of dolls, line drawings, and props.

3. There are reliable age effects in children's suggestibility, with preschoolers (the age of the Perkins children) being most vulnerable to a host of factors that contribute to unreliable reports.

4. Errors that children make as a result of suggestive techniques include not only peripheral details, but also central events that involve their own bodies. At times children's false reports can be tinged with sexual connotations. In research studies, young children have made false claims about "silly events" that involved body contact (e.g., Did the nurse lick your knee? Did she blow in your ear?), and these false claims persisted in repeated interviewing over a three-month period (Ornstein, Gordon, & Larus, 1992). Young children falsely reported that a man put something "yuckie" in their mouth (Poole & Lindsay, 1995,2000). Preschoolers falsely alleged that their pediatrician had inserted a finger or a stick into their genitals (Bruck, Ceci, Francoeur & Renick, 1995) or that some man touched their friends, kissed their friends on the lips, and removed some of the children's clothes (Lepore & Sesco, 1994). A significant number of preschool children falsely reported that someone touched their private parts, kissed them, and hugged them. (Goodman, Rudy, Bottoms & Aman, 1990; Goodman, Bottoms, Schwartz, Kenney & Rudy, 1991; Rawls, 1996; Bruck, Melnyk & Ceci, in press). In addition, when suggestively interviewed, children will make false allegations about nonsexual events that could have serious legal consequences were they to occur. For example, preschoolers claimed to have seen a thief in their day care (Bruck , Ceci & Hembrooke, 1997).

2. The "mix" of suggestive interviewing techniques in conjunction with the degree of interviewer bias can account for variations in suggestibility estimates across and within studies. If a (biased) interviewer uses more than one suggestive technique there is a greater chance for taint than if he uses just one technique.

6. Even though suggestibility effects may be robust, the effects are not universal. Thus, even in studies with pronounced suggestibility effects, there are always some children who are highly resistant to suggestion; and there are also some adults who are highly suggestible. Further, 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., Garven et al., 1997; Thompson et al.,1997).

Despite the variation in suggestibility, social scientists have not identified a particular cognitive, personality, or temperament style that can predict on an individual level which child will or will not fall sway to suggestive interview techniques. To date the strongest determinant is age.

7. Finally, although I have concentrated on the conditions that can compromise reliable reporting, it is also important to acknowledge that a large number of studies show that children are capable of providing accurate, detailed, and useful information about actual events, some of which are traumatic (e.g., see Fivush, 1993; Goodman, Batterman-Faunce & Kenney, 1992 for a review). What characterizes these studies is the neutral tone of the interviewer, the limited use of misleading questions (for the most part, if suggestions are used, they are limited to a single occasion), and the absence of any motive for the child to make a false report. When such conditions are present, it is a common (although not universal) finding that children are much more immune to suggestive influences, particularly about sexual details. When such conditions exist in actual forensic or therapeutic interviews, one can have greater confidence in the reliability of children's allegations. Unfortunately these conditions were not present in the Perkins case.

Suggestive Interviewing in Perkins: Detailed Examples

The pattern of disclosure was similar for all three children. At first the children were silent, they made no allegations of abuse. When first explicitly asked about abuse, they denied. With repeated questioning and with repeated interviewing, the children slowly began to disclose. Over time, the children provided additional details in their narratives and often dropped out earlier ones. This pattern by itself does not necessarily force the conclusion that the children's reports were unreliable. However, in light of the fact that the allegations emerged in interviews that contained a number of suggestive components (e.g. questions were repeated, leading information was included, anatomically detailed dolls were used, peer pressure, selective reinforcement and a range of emotional atmospherics were included), then there is a very high risk that the children's reports were unreliable. I provide examples of the suggestive questioning for each of the three children.

JACLYN

Trish Perkins questioned her four-year-old daughter about Bruce at the beginning of September, 1992. Although Jaclyn had never made any mention of sexually inappropriate touching, Trish asked her, "Has Papa Bruce touched you in a way that you don't like? Has he touched your private parts?" (Bynum, tape transcript p. 2). Jaclyn denied touching. Trish testified that she had repeated the question two or three times that evening (Tr 711). Finally, in a very "silly way", Jaclyn assented (Tr. 711) and provided a "general answer, "he touches me"" (Bynum tape transcript, p.2).

Then either the same day or a few days later Trish tried to find out if Jaclyn had been threatened. Trish asked Jaclyn a series of leading questions which eventuated in an ambiguous report of a threat. (The questions were asked while the mother and child are playing with letters and their sounds.)

> "I said you know, what did Paw-Paw Bruce say to you, did he say to you don't tell and she said she said" ssss, guess Mommy". So I started guessing every S word that I can think of without trying to be too, ya know. I said "suck" and she said no and I was relieved and I said, "six", "secret", no, no. (and she said) "Sorry, Bruce said he was sorry." (Bynum Transcribed interview p. 6).

A day or so later, for the first time, Larin talked to Jaclyn. The following is Trish's summary of the conversation (Bynum interview transcript, p. 6).

> Larin: Mommy says that Paw-Paw Bruce touched you and that you told Mommy that he touched you
> Jaclyn: No (acting very silly)
> Larin: Well Mommy said that Paw-Paw Bruce said he was sorry, well why did he say he was sorry ?
> Jaclyn: Because he touched me
>
> *COMMENT: Jaclyn's response was consistent with the wording of the previous leading question, Bruce touched you.*
>
> Larin: Where did he touch you?
> Jaclyn: (pointed to arm).
> Larin : Where else?
> Jaclyn: (pointed to shoulder).
> Larin : Where else?
> Jaclyn: (pointed to leg).
> Larin : Where else?
> Jaclyn: (pointed to genitals).
>
> *Comment: It is possible that Jaclyn did not want to talk about sexual touching because of embarrassment and fear, and that she needed a supportive and leading environment to extract reports of abuse. But given the scientific literature there is also the strong possibility that false claims of sexual touching emerged as a result of repeated and leading questions. Jaclyn finally provided her parents with the answers they wanted to hear rather than with accurate reports of her own experiences.*

Later the same night, after much more questioning, the parents brought out dolls for the child to tell more of her story of abuse (According to Trish, Dr, Kammholz had used dolls to question the children, so this was not the child's first experience in demonstrating sexual actions on props). In response to her mother's direct questions, Jaclyn demonstrated oral sex with her brother and then with her grandfather (Bynum transcript p. 9, 12). Throughout this part of the questioning, Jaclyn was told she was brave, that her parents were proud of her and that they would protect her. These statements created the climate (emotional atmospherics) of encouraging the child to talk about sexual abuse (true or false).

Trish Perkins never denied the fact that her daughter's statements were coerced. She recognized and seemed proud of the fact that she was able to elicit disclosures through highly suggestive techniques (Tr, 735). According to Joyce Pruitt, Trish told her on the phone, that she was a professional RN and she had dug the truth out of Jaclyn but Josh hadn't said anything yet but he will."(Tr 2058)

A few days after the first disclosure, Jaclyn was interviewed by Bynum. There are several noteworthy suggestive components in the Bynum interview. First, Jaclyn's mother, Trish, was present. Despite Bynum's statement at trial that this was important to make the child feel at ease (Tr 1201-1202), this procedure tainted the evidence. Trish intruded a number of times during the interview, prompting the child to provide specific answers and sometimes providing the child with the answers. For example,

> BYNUM: Oh, what do you call it, do you call that? a vagina, you do? Ok, well I have heard that word before, I am not embarrassed to hear it, what do you call this back here? What do you call that?
> **TRISH:** Bottom, hinney hole
> BYNUM: Do you call it bottom or hinney hole?
> JACLYN: Bottom
> (p. 4 of transcribed interview).

Furthermore, even had Trish remained silent, the physical presence of the mother is a potential taint factor. The point of an investigative interview is to determine what the child says happened. With her mother present, the risk increases that Jaclyn will merely repeat the details she provided the mother whether or not these truly happened.

A second suggestive aspect of Bynum's interview is the almost sole reliance on specific leading questions that contain information previously supplied by the parents. Jaclyn was not asked open-ended questions, she was not given the opportunity to say in her own words what happened when she visited her grandfather..

A third suggestive aspect is that Jackie was asked a numerous leading questions such as:

-Who touched your vagina?
- Did you ever see a penis before (who?)
- Who was lying down on a bed?
-Has anyone touched your vagina with the mouth?
-Did you put your mouth on a penis?
-Who sucked Josh's penis?

The interesting aspect of her reports was that she usually provided an affirmative answer or supplied a name, but she never initially named her grandfather as an abuser for any of these highly leading questions. It is only with several repeated questions ("who else?") that she named Bruce.

> BYNUM: Who has touched you there (Bynum is pointing to the genitals and buttocks of an anatomically detailed doll)
> JACLYN: Mandy
> BYNUM: ...Has anybody else touched you there?
> JACLYN: Craig
> BYNUM: Anybody else?
> JACLYN: Un-un (No)
> BYNUM: : Has anybody ever touched you right here?
> JACLYN: Um-huh

BYNUM : Who?
JACLYN: Mandy, Craig and Jennifer
BYNUM : Mandy, Craig and Jennifer, no one else has ever touched you?
JACLYN: Un-un

BYNUM: Has there ever been any big person there when they were touching you?
JACLYN: Uh-huh
BYNUM: Who has been there?
JACLYN: Jennifer
BYNUM: Jennifer is the big person you know -
JACLYN: Uh-huh
BYNUM: Uh-huh, nobody else, no big man, huh, OK, well has anybody ever been, a big person been near, when you have been touched here?
JACLYN: Uh-huh
BYNUM: Who?
JACLYN: My brother that is 6
BYNUM : Your brother that is 6, what is his name?
JACLYN: : Josh

BYNUM: Has there been a great big person like me ever touched you there, no body has, you have never been touched by any body like that before, have you ever told any body that a big person has touched you there before, huh, never
JACLYN: Just my mom and my dad
BYNUM: Just your momma and your dad
JACLYN: Um-huh

BYNUM: OK, have you ever seen anything like this on a man before, on a boy, who, what do you call that
JACLYN: Penis
BYNUM : Penis, you did good, that is great, have you ever seen, whose penis have you seen before
JACLYN: My brother's
BYNUM : Have you ever seen a big person's penis before
JACLYN: Um-huh
BYNUM : Who?
JACLYN: My Paw-Paw

Thus in the initial investigatory interview, Jaclyn made no spontaneous claims of abuse involving her grandfather who was named only as a result of very suggestive and repeated questioning in the presence of her mother.

Soon after, Jaclyn began therapy with MRM who did not learn any details from the child for several sessions. MRM testified that she used directive (i.e., suggestive) techniques with the child whenever she needed help to bring up topics (Tr. 984). For example, MRM testified:

> "She chose to play with the stuffed animal that day and she was being a veterinarian and I intervened by just pointing out the difference in what doctors do and what doctoring is and she seemed to understand there was a difference between that. (Tr. 1019)
> *Comment Jaclyn's innocent play with a stuffed animal was interrupted by MRM who inserted the topic of sexual touching and explicit sexual details.*

Dolls were used through out the treatment sessions to assist Jaclyn to talk and think about sexual abuse. Thus, Jaclyn's sexual knowledge may have emerged as a result of repeated requests to show sexual abuse with the dolls.

Finally, throughout the therapy sessions, Jaclyn provided details of abuse that were not disclosed previously to her mother, to Bynum or even to Kammholz. The following table provides examples of how Jaclyn's allegations changed or evolved over time.

Initial Allegations to Mom/Kammholz /Bynum	Allegations made during Therapy
Bruce said he was sorry	Bruce said, I will throw you off a pole, feed you to spiders and to snakes
Games played with Mandy and Craig where toys were put in tiney holes (p. 456)	Bruce put a bottle in her rectum Bruce put fingers in the cat's rectum Bruce tied us up
Grandma Worms was there	Carol was there
Oral copulation by Jennifer and Josh	Oral copulation by grandparents

Although some of the same themes co-occurred in different interviewing contexts, the abusers and the intensity of the accusations changed. As explained above, when new details enter children's testimony after repeated and suggestive interviews, there is a high probability that the new details are false.

Jaclyn's pattern of disclosure is characterized as one in which details were drawn out of the child by means of suggestive and repeated interviews. The child lived in an atmosphere where sexual abuse was suspected from the time she was at least 3-years old. She reported that her grandfather had abused her only after her mother had directly confronted her and questioned her many times. Jaclyn tried to tell her father that abuse had not occurred, but with enough repeated questions, she produced statements that were consistent with the belief of abuse. These statements were elaborated and changed over time. Based on the scientific literature, it is my expert opinion that the interviewing and questioning techniques used with Jaclyn have a higher than usual risk of eliciting false disclosures of abuse especially when used over such a prolonged period with a child as young as Jaclyn.

MANDY

Prior to September 8, Mandy had made no allegations involving her grandfather. Based on the allegations made by Jaclyn, Patty, first asked Mandy a very suggestive question, "You know what Jaclyn told her mom that Papa Bruce had touched. her and touched her in a way that she doesn't like to be touched", Mandy denied (She acted like her mother was crazy). When Patty repeated the subject, "Well, why would Jaclyn say something like that?", Mandy had no answers. The next day, Patty continued the questions, "Mandy if a big person has touched you there, I need you to tell me so that I can protect you.". Mandy denied that she had seen Bruce abusing Jaclyn or Josh. Only after assuring

her daughter that she would not be angry did Mandy claim that Bruce had touched her on the bottom with a finger. Mandy then demonstrated the alleged abuse with two dolls. According to Patty she did not understand much of what Mandy told her. Mandy claimed that Bruce took all kinds of food and wiped it on her (she pointed to her bottom). Then Mandy pointed to the genitals in a picture that she had drawn and claimed that Bruce would eat the food. Throughout this session Mandy was laughing. (See Patty/Bynum interview, transcript of audiotape pp 1-4).

Bynum interviewed Mandy and her parents. Although the father was asked to leave the room, the mother was allowed to remain while Bynum questioned Mandy. There are several noteworthy aspects of this highly suggestive interview.

First, Patty was highly intrusive, prompting the child at a number of different occasions.

> BYNUM: What do you call this right there?
> MANDY: Un-un
> BYNUM: You don't know, you don't have a name for that, do you ever call that Sarah
> PATTY: : Say vagina
> MANDY: Vagina
>
> BYNUM: Did he ever put anything on your vagina, what did he put on your vagina?
> PATTY: A toy
> BYNUM: He put what?
> MANDY: A toy (Transcript of Mandy Bynum interview p. 4)

Thus Patty actively supplied the required information. As a result it is not known if Mandy's answer reflected her own direct experiences and knowledge or that of her mother's.

Second, Mandy's allegations of abuse were responses to specific leading questions in a doll-centered interview. As discussed above, the scientific literature shows that the combination of these factors (asking the child leading questions while manipulating the doll) can have devastating effects on the accuracy of her report.

Mandy was then treated by Nelke for childhood sexual abuse. As was the case with Jaclyn, the claims of abuse escalated and became more elaborate over time. The chart details the types of allegations made to the mother, Bynum ,Nelke, the prosecutor (direct testimony) and the defense attorney (cross examination) It should be noted that from the record it appears that most of these were obtained through leading questions.

As the next table shows, Mandy's testimony is marked by its inconsistency. She changed her answers to questions across interviews and within interviews. She recanted and denied previous testimony in her cross-examination, and in some cases in her direct examination. A reading of the trial transcripts clearly shows that Mandy was harassed by questions that she often did not want to answer. She constantly stated that she was tired and that she did not remember.

As was the case for Jaclyn, my expert testimony is that the suggestive conditions under which Mandy were repeatedly interviewed (even through trial) render her testimony unreliable.

[redacted] [redacted] [redacted] C[redacted]ster[redacted]of N[redacted]dy'[redacted]cus[redacted]ns [redacted]n S[redacted]mb[redacted]th to[redacted]oss-[redacted]mi[redacted]on [redacted] P[redacted] [redacted] [redacted] [redacted]

Mother's Report/ Sept 8	Bynum on Sept 8	During Therapy	Direct Examination	Cross Examination
Bruce touched bottom with finger	Bruce touched her vagina	Bruce touched her vagina	Touched inside of bottom and vagina with finger	
	Never saw him touch Jaclyn	Bruce put penis in Mandy's mouth & bottom	Penis in Mandy and Jaclyn's mouth & bottom Touched Jaclyn everywhere	Did not touch Jaclyn
	Put a toy in vagina—lego Didn't stick anything else		Penis in vagina	
	It hurt			
If tells, mom will punish her—put her in her room		Don't tell because Bruce will get in trouble Will chop Mandy into a million pieces	Chop up and feed to chicks and horses (72) DOESN"T KNOW WHY HE SAID THAT (73)	
Happened in his bedroom Bruce covered me up	They were in bed			
	Jaclyn was there	Children & baby Jillian locked in the closet.		
		Dog claws and tongue on her bottom		
Joshua touched Bruce's penis	She never saw a penis	Saw Bruce's penis	Touched Joshua's penis and bottom with hands and penis -Recants (70)	Bruce did not touch Joshua's penis (T.133)
	Never touched a penis			
Bruce wiped ketchup & mustard on bottom. Ate it	Put ketchup on vagina	MAY 24: Sour cream & catsup on bottom, JUNE 7: Sour cream came from penis. Put it on bottom and wiped it off Thinks Ketchup came from penis	Sour cream came out of penis. Rubbed it on our bottom	
	Jennifer, Bruce and Jacob touched her vagina			Jennifer & Jacob present Recants (120)
		Carol touched her vagina		
		Bruce and Carol stuck tongue in her mouth		
			Bruce took off Jac, Josh Mandy's clothes (60) RECANTS (70)	DENIES (122)
			No pornographic pictures (72)	

JOSHUA

Joshua Perkins, the oldest of the alleged child victims, was the most resistant to the suggestive interviewing that occupied his life from at least March 1992 until the time at trial.

--Joshua resisted talking to Carolyn Kammholz (387).

--Joshua did not disclose any information to *MRM* who saw him for several sessions (Tr, 1035).

--Joshua then was treated by Dr. RM. Despite the fact that he has heard his parents reminding RM about their concerns of abuse, Joshua still did not disclose (Tr. 125).

-- According to Joshua's parents, he said he would be killed by his grandfather if he ever told of the abuse. But even by the time of trial, he had not disclosed any specific threats to his therapist who interpreted his silence as "shutting down" --Although RM testified that Joshua began to disclose abuse, there were no details in the record about what was disclosed.

Thus Joshua consistently denied abuse despite the persistent pressure of many of his interviewers to disclose. And yet, after this long period of silence, Joshua testified at Bruce Perkins' trial. Because of the number of repeated interviews and the long strings of denials, there is a real concern that Joshua's monosyllabic answers of "yes" and "no" to leading questions at trial were not reliable.

The following chart shows the questions assented to and denied during his trial testimony. (*Note.* In most cases Joshua did not produce the follow statements, but rather answered yes or no to questions phrased from these statements (e.g., Did you go upstairs? Yes).

Assents to leading questions	**Denials** to leading questions
Went upstairs	
Jaclyn was there	
Saw Bruce do something to Mandy Touched Mandy's genitals but not bottom	Mandy was not there
Bruce touched Jaclyn 's genitals and buttocks with his hand	Didn't do anything else Didn't put anything inside her mouth
She had her clothes off	
Bruce's clothes were off	
Saw Bruce's penis	Nothing came out. No sour cream
Bruce touched his penis	Never touched anyone with his penis Bruce didn't touch Josh Josh didn't touch Bruce's penis
On another day took off clothes	Josh didn't take off clothes Never took off clothes playing with his sister.
Went in a closet	
He scared me Said he would kill me	He didn't scare me
Cut me up into a million pieces	Don't know how he would kill me
Carol took pictures of naked children	Bruce didn't take pictures Never saw pictures of naked people
Played doctor with Craig Shannon Bruce taught how to play doctor	
	No oral copulation

After almost two years of relentless questioning of the type most likely to create false reports, Joshua produced a few accusations some of which were either not corroborated by his own testimony or by his sister's or cousin's testimony, This pattern could be indicative of a child who has been coached or suggestively interviewed and if this is indeed the case Joshua's testimony was irreparably tainted.

IV. THE SCIENTIFICALLY UNGROUNDED AND INACCURATE TESTIMONY OF THE STATE'S EXPERT WITNESSES

Credibility

It has often been stated that it is easy to detect false reports that are the result of suggestion, because children are merely "parroting" the words of their interrogators. However, evidence from the past decade, provides no support for this assertion. First, we have found that children's false reports are not simple reflections or monosyllabic responses to leading questions. Under some conditions, their reports become spontaneous and elaborate, going beyond the suggestions provided by their interviewers. For example, in the Bruck, et al., study (1997), children's false reports contained the prior suggestion that they had seen a thief take food from their day care; but the reports also contained nonsuggested details such as chasing, hitting, and shooting the thief (also see Bruck, Ceci, Barr & Francoeur, 1995). Second, subjective ratings of children's reports after suggestive interviewing reveal that these children appear highly credible to trained professionals in the fields of child development, mental health, and forensics (e.g., Leichtman & Ceci, 1995, Ceci et al. 1994); these professionals cannot reliably discriminate between children whose reports are accurate from those whose reports are inaccurate as the result of suggestive interviewing techniques. The children who provided the false reports spoke sincerely and provided accounts laden with emotion and perceptual details.

Third, results of a recent study revealed that linguistic markers do not consistently differentiate true from false narratives that emerge as a result of repeated suggestive interviews. In the Bruck et al., study (1997) where children were repeatedly and suggestively interviewed about true and false events (described above), the children's narratives of the false events actually contained more embellishments (including descriptions and emotional terms) and details than their narratives of the true events. Also the false narratives had more spontaneous statements than the true narratives. Although for the most part, the details in false stories were realistic, as suggestive interviews continued, children inserted bizarre details into their stories.

Another line of study suggests that even well-trained professionals cannot reliably differentiate between true and false reports of sexual abuse (Realmuto, Jensen and Wescoe, 1990), even when provided with extensive background information about the case (Horner, Guyer & Kalter, 1993a, 1993b).

Given the scientific background, the prosecutions' experts provided the jury with inaccurate testimony regarding their ability to detect deception and truth in the statements of the children. Their explanations were not only unscientific but inaccurate. The following are a few of the numerous inaccurate examples of their testimony:

When RM was asked about bizarre details of a child's story, such as a dog's penis being cut off, he replied that its significance was that it was true, that those events actually happen and that such events are so far outside the child's experience that they have to be true.(Tr. 2605).

Comment. This is a subjective judgment that is based on his belief that similar accusations made in previous cases were also true. But all could be the result of suggestive interviewing procedures. Also, in the present case, this bizarre allegation did not survive into the courtroom.

MRM testified that because Jaclyn's play was very spontaneous she was not coached or suggested (1022).

MRM was certain that Trish was not suggestively interviewing or coaching Jaclyn because of how difficult it was for Jaclyn to talk to MRM and because of her reactions in the therapy session (1146).

Comment: Trish admitted to her coercive interviewing techniques.

CN testified, "When children have been coached, when they talk about the sexual abuse, they talk about it in a different way than when its genuine. Often there is no affect involved, they're not showing an kind of emotional reaction to what they're talking about. Children who are coached get a lot of outside reinforcement for making statements about sexual abuse. For instance, when they are going to see a therapist, what you oftentimes see is they will come into the room and make their statements: "Yeah, you know, this person sexually abused me". Or: "This person touched me on my private parts". It tends to be very general and vague in terms of the allegations they are making. And then they say it. And they don't show any distress or anything. They almost look at the professional for some praise for making the statement that they made and then it doesn't seem to have an effect on them and then they go on to play activities or to talk about different topics. When children can talk about details it gives more validity to their allegations. (Tr. 819).

Comment: The scientific literature does not support any one of CN's opinions. In fact, it often shows just the opposite.

Inaccurate Testimony about the Suggestibility Literature

The therapists either misinterpreted or else gave a scientific veneer to many of their opinions. Below is a very partial list of these statements.

RM testified that children four years of age and older were not more suggestible than adults under certain circumstances. (Tr. 2597.)

Based on thorough reviews of the published scientific literature, this statement is false. In most studies, there are developmental differences between children of all ages and adults, and these differences are most prominent for children under the age of 7-years (e.g., Ceci & Bruck, 1995).

RM testified that children under three are suggestible and their susceptibility to suggestion depends on how comfortable they are under the circumstances and how well they know what it is they are being asked about If the subject is unclear to them, they are more suggestible (Tr. 2598).

RM's opinion implies that there is a wealth of data on this subject. There is not. We know that three-year olds (and four year olds and five year olds etc.) are suggestible in a range of situations. Sometimes accuracy is increased when children are provided highly situations, at other times it decreases (Geiselman, et al., 1990). Furthermore, the logical implication of his statement about the relationship between knowledge and suggestibility is that it is easier to suggest sexual abuse (an unfamiliar subject to preschool children) than a more familiar subject to a young child. This is certainly a reasonable hypothesis that he does not apply to the present case. Namely, the Perkins children only assented to suggestive questions because they did not understand their meaning.

RM testified that the research that had been done in looking at errors in children's report suggest that "children are extremely likely to omit details from their disclosures and that it's extremely rare that children include details that didn't actually happen." (Tr. 2598)

As support for this statement RM described a study by Saywitz et al (1991) that included girls ages 5 and 7. When asked about a recent medical examination with the use of anatomically detailed dolls, children in this study were more likely to omit details of a genital examination than to make false claims of genital touching. RM further stated that there were no false claims of genital touching. The latter statement is incorrect. The study in fact, demonstrated that 2.86 percent of the children studied falsely claimed genital touching when in fact, it had not occurred; 5.56 percent falsely reported anal touching when it had not occurred. Nevertheless, RM did get the gist of the study, however, he did not accurately contextualize the results for the court. The important aspect of this study was that (a) children were asked about touching to the genitals and the buttocks by a neutral investigator and they were only asked two direct questions in one interview about genital touching (Did the doctor touch your genitals? And (pointing to the genital area of the doll—"Did the doctor touch you here?). When children had been touched they correctly assented in 86% of the cases after the second question. Thus although they may have been likely to deny the first time round, when asked again within the same interview, most did provide the correct response. Compare this to the conditions under which the Perkins children were questioned. It is not only that the Perkins children denied touching when asked the first time, but even when pressured after many questions, they continued to deny that there had been any touching. Based on the scientific literature, this continued pattern of denial should raise concerns that perhaps the denials were accurate and that there was no inappropriate touching.

Apart from the Saywitz et al., study that does show that upon first questioning children are more likely to omit details than to insert wrong details, there are numerous studies in the literature that show that when children are interviewed more than once, or when the interviews are quite suggestive, then children commonly include details that never happened. (e.g., Bruck et al, 1997; Thompson et al., 1998; Poole & Lindsay, 1995; Leichtman & Ceci, 1995) even when they are asked about genital touching (e.g., Bruck et al., 1995; Bruck et al., in press).

RM misinformed the jury about the important details of this literature but he also failed to qualify his testimony in terms of the implications of the study for the Perkins case.

RM testified "The research on children's reports suggests that kids report pretty accurately whether a parent is there or not, but they report more information which is also accurate if they are comfortable in a situation. And one of the conditions of comfort sometimes is having a parent present (Tr. 2604).

I have conducted several thorough reviews of the scientific literature over the past decade. I know of no studies that have been carried out on the effect of the presence of the caretaker on the accuracy of children's reports when these are elicited in suggestive contexts. Furthermore to make the unqualified statement that children report accurately, is misguided. When children are suggestively interviewed their reports can be quite inaccurate.

The Therapists Failed to Understand the Concept and Influence of Suggestive and Biased Interviews

According to the expert witnesses, the children must have been abused because their statements *contained* details that could only be expressed if the children had had direct experience. The therapists discounted the possibility that the children's statements were the result of "coaching". However, coaching should not be equated with suggestive interviewing. Coaching occurs when a child is deliberately told or rehearsed to make certain statements. Suggestion however is more implicit and not necessarily deliberate. The experts did not seem to understand that suggestions could result in very credible but very wrong stories, and that some children appear so credible because they come to believe that the suggestions actually happened

The experts also did not consider the role of contamination in the present case. Contamination is one powerful source of suggestion. In the present case the children on numerous occasions heard their parents providing therapists or investigators with details of the case. When children hear details of sexual activities, this not only provides the children with the concepts but also with the linguistic labels. For example, while the children were present, Trish Perkins (and perhaps Patty Perkins) told Kammholz about the details of a number of sexual activities (that may or may not have occurred) (TR. 411). At the police station, the parents related the details of their children's disclosures as well as other concerns about the children's grandparents. Although Bynum denied that the children could overhear these interviews, it is clear that the children were interacting with their parents during these interviews

and were within earshot of these conversations (Bynum Jaclyn transcript p. 4; Mandy Bynum transcript, p. 2 and 3). Also, in the therapy sessions with RM, Josh would be present as his parent (s) told about their concerns of the case. As Josh heard these he became more and more anxious (Tr. 125. 133). Thus the children had numerous opportunities to listen to their parents talk about sexual details. And with suggestive interviewing, it is the details of these overheard conversations that they may have reported rather than any direct experience.

Another example of contamination concerns the children's own interactions with each other. For example, Kammholz interviewed the three children together. Although there were no allegations about Bruce in these sessions, this still produced the opportunity for the children to be bathed in each others stories. When pressured to produced details of abuse, the children may have produced stories that evolved in these therapy sessions.

It is also possible that the parents communicated the allegations of one child to another and the other child then incorporated these into their report. For example, on May 24th Mandy Perkins drew a picture of her grandfather , the children and a jar of sour cream. In the next session, on June 7th Mandy's mother told the therapist that Mandy had admitted that she had made a mistake and that she really meant to say that the sour cream came from his penis. Two days later, on June 9th, MRM, Jaclyn's therapist reported the following in her therapy session:

> I then asked her if she saw her Grandpa Bruce's penis; and she said, "Yes."... I asked her what he did with his penis; and she said, "He put it in my mouth." And I said, "Then what happened?" And she said, "Mandy called it sour cream, but it was pee pee." (Tr. 1027)

This example clearly shows that the children's mothers knew what each child was saying in therapy and that they communicated this to their own child. The fact that Jaclyn reported the same incident with sour cream within two days of her cousin and she knew exactly what her cousin had said and what her cousin had meant to say reveals how some of the children's statements were the result of contamination and thus were tainted. Although this and previous examples do not involve "coaching" (the only form of taint that the therapists' considered and rejected), contamination among the child witnesses and their families can serve as an important source of taint and could explain the emergence of some of the children's disclosures.

Another piece of evidence to counter the therapists' argument that the children's reports must reflect their own experiences because how else would children learn about such matters comes from an analysis of the Bynum interviews. When Bynum interviewed Jaclyn and Mandy, they made no spontaneous claims of abuse; rather, these were generated by their mothers and through Bynum's questions which were filled with references about oral copulation, and body parts. These types of

questions could influence the child's knowledge base and could produce statements that contain sexual information even though the statements have never been personally experienced.

The major point is that that the Perkins children were constantly asked about abuse and asked to demonstrate abuse (Kammholz used dolls, the parents used dolls etc., the therapists showed the children sexually graphic books) for a period of time. It is these situations that can produce detailed elaborations of events that did not happen even for sexual events. All these factors in combination with the suggestive and unrelenting questioning techniques that focused on sexual abuse could easily result in the production of false and elaborate narratives of sexual abuse by children the ages of those in Perkins.

V. CONCLUSIONS

I was asked to review the materials in Perkins that related to the interviewing of the children. Based on my review of these materials I found that:

--There was a long history of suspicion of sexual abuse in the Perkins families.
--The young children were questioned about abuse from a very young age.
--The children initially never made any spontaneous allegations about their grandfather.
--The children denied that their grandfather had abused them when they were first asked by their parents.
--The children began to make disclosures only after repeated questions and interviews. The questions were leading (containing suggestions of sexual abuse). The children were asked to demonstrate abuse with dolls. The children were encouraged through a number of atmospheric and selective reinforcement techniques (you are so brave, Mommy will protect you) to disclose abuse. The children were also told what their siblings or cousins had purportedly said in an attempt to get them to disclose. Children were interviewed together and often they overheard their parents' providing details to the police and therapists.
-Finally, the degree of bias was high among the interviewers. The therapists' accepted the parents' word that the children were abused and interpreted all their behavior (including denials) as symptomatic of abuse. The police officer who first questioned the children structured the interviews to elicit statements from the children that would confirm the parents' initial reports and suspicions.

In light of this information and based on the scientific literature, it is my expert opinion that there was sufficient bias and suggestive interviewing of the young Perkins children to taint their testimony. There is a very high risk that their allegations of abuse were unreliable and merely a reflection of the misleading suggestions provided by their interviewers.

I declare under penalty of perjury that the foregoing is true and correct to the best of my knowledge and belief.

Executed on May 15, 2000, at Baltimore, MD.

Maggie Bruck

Maggie Bruck, Ph.D

Index of Materials Reviewed

I have reviewed the following materials before reaching my conclusions for this affidavit:

The trial testimonies of:

- Nelke
- Kammholz
- Richardson
- McLaughlin
- Detective Bynum
- Mandy Perkins
- Joshua Perkins
- Trish Perkins
- Paula Porter
- Joyce Pruitt

Audiotapes or transcriptions of audiotapes of Bynum's September interviews with the Perkins families and their children. These include interviews with

- Patty and Lann Perkins
- Trish and Larin Perkins
- Mandy Perkins
- Jaclyn Perkins

Notes from Dr. Berliner's interview with Patricia Perkins.

Psychological Bulletin
1993, Vol. 113, No. 3, 403-439

0033-2909/93/$3.00

Suggestibility of the Child Witness: A Historical Review and Synthesis

Stephen J. Ceci and Maggie Bruck

The field of children's testimony is in turmoil, but a resolution to seemingly intractable debates now appears attainable. In this review, we place the current disagreement in historical context and describe psychological and legal views of child witnesses held by scholars since the turn of the 20th century. Although there has been consistent interest in children's suggestibility over the past century, the past 15 years have been the most active in terms of the number of published studies and novel theorizing about the causal mechanisms that underpin the observed findings. A synthesis of this research posits three "families" of factors—cognitive, social, and biological—that must be considered if one is to understand seemingly contradictory interpretations of the findings. We conclude that there are reliable age differences in suggestibility but that even very young children are capable of recalling much that is forensically relevant. Findings are discussed in terms of the role of expert witnesses.

Since the turn of the century, psycholegal scholars have examined the suggestibility of children's testimony in an effort to determine whether they would be credible witnesses. A major issue in this research concerns the degree to which heightened levels of suggestibility may affect children's ability to accurately report what they have witnessed.

In this article, we review and integrate the entire corpus of 20th-century social science research concerning young children's presumed suggestibility. In the past 10 years, more research has been conducted on the suggestibility of child witnesses than in all of the prior decades combined. This increased research has been motivated by practical concerns: Young children are increasingly being called to testify in court, particularly in sexual abuse cases. Because the earlier literature was criticized for its lack of methodological sophistication and poor ecological validity, it was deemed unsatisfactory for addressing the issue of children's testimonial competence. However, as we show, although contemporary cognitive, social, and developmental psychologists have attempted to provide insights into the intricacies of children's testimonial competence in ecologically relevant settings, the literature is riddled with contradictory interpretations of results. On the one hand, children are described as highly resistant to suggestion, as unlikely to lie, and as reliable as adult witnesses about acts perpetrated on their own bodies (e.g., Berliner, 1985; Goodman, Rudy, Bottoms, & Aman, 1990; Jones & McGraw, 1987). On the other hand, children are described as having difficulty distinguishing reality from fantasy, as being susceptible to coaching by powerful authority figures, and therefore as potentially being less reliable than adults (e.g., Feher, 1988; Gardner, 1989; Schuman, 1986; Underwager & Wakefield, 1990).[1]

The purpose of this review is to provide a historical integration of the research in this area. We attempt to show how the research has reflected cultural, legal, and psychological concerns of the day. Although our review shows that there is still controversy regarding some aspects of children's suggestibility, we try to reconcile this controversy by taking issue with extreme views regarding children's competence. We argue that although there is controversy, it is less the result of inconsistent data than of how these data are interpreted. To resolve this apparent controversy, we reorient this debate to one concerning the causal mechanisms underlying suggestibility in order to understand under what conditions children are or are not more suggestible than adults.

We begin this review by describing two recent court cases in which child witnesses provided critical eyewitness testimony. These cases serve as "windows" through which the points we make later can be viewed; namely, how accurate are children's recollections of everyday events? How suggestible is the child witness? How much difficulty does the child have distinguishing reality from fantasy? How honest are children?

The Wee Care Nursery Case

Margaret Kelly Michaels, a 26-year-old nursery school teacher, was accused of sexually abusing children at the Wee

Stephen J. Ceci, Department of Human Development, Cornell University; Maggie Bruck, Department of Psychology, McGill University.

This research article was supported by a National Health and Welfare Scholar Award to Maggie Bruck, by Natural Sciences and Engineering Research Council Grant 0GP000 A1181 to Maggie Bruck, and by National Institutes of Child Health and Human Development Grant ROI HD 25775 to Stephen J. Ceci.

We are particularly indebted to the thorough, thoughtful, and helpful comments provided by the four reviewers of this article, Stephen Lindsay, Peter Ornstein, Amye Warren, and Maria Zaragoza—as well as by the comments and advice of other colleagues: Chuck Brainerd, Beth Loftus, John Doris, Steve Golding, and Urie Bronfenbrenner.

Correspondence concerning this article should be addressed to Stephen J. Ceci, Department of Human Development, Cornell University, Ithaca, New York 14853.

[1] We do not mean to imply that proponents of these opposing positions have wholeheartedly endorsed extreme views of the child witness because both camps express the belief that children are capable of high levels of accuracy, provided that adults who have access to them do not attempt to bias their reports.

Care Nursery School. She was said to have licked peanut butter off children's genitals; played the piano while nude; made children drink her urine and eat her feces; and raped and assaulted them with knives, forks, spoons, and Lego blocks. She was accused of performing these acts during school hours over a period of 7 months. No alleged act was noticed by staff or reported by children to their parents. No parent noticed signs of strange behavior or genital soreness in their children or smelled urine or feces on them.

The first suspicion that Kelly Michaels abused her charges occurred 4 days after she had left the Wee Care Nursery School. A 4-year-old former student of Kelly's was having his temperature taken rectally at his pediatrician's office and said to the nurse, "That's what my teacher does to me at school." When asked to explain, he replied, "Her takes my temperature" (Manshel, 1990, p. 8). On the advice of the pediatrician, the child's mother notified the state's child protective agency. Two days later, when the child was interviewed by the assistant prosecutor, he inserted his finger into the rectum of an anatomical doll and reported that two other boys also had their temperatures taken. When questioned later, these two boys denied the claim, but one indicated that Kelly Michaels had touched his penis. The first child's mother then told a parent, who was a board member, of her son's disclosures. This board member interrogated his son about Kelly Michaels touching him inappropriately, remarking that "he was his best friend and that he could tell him anything" (Manshel, 1990, p. 126). His son said that Kelly had touched his penis with a spoon. The Wee Care Nursery School sent out a letter to parents, informing them of an investigation of a former employee "regarding serious allegations." In a subsequent meeting, a social worker explained to the parents that sexual abuse of children is very common, with one out of three children being victims of an "inappropriate sexual experience" by the age of 18 years. She encouraged parents to examine their children for genital soreness, nightmares, bedwetting, masturbation, or any noticeable changes in behavior and to have them examined by pediatricians for injury. Soon, there were many more allegations against Kelly Michaels. Two and one half years later, she was convicted of 115 counts of sexual abuse against twenty 3- to 5-year-old children. She is serving a 47-year sentence.

The Country Walk Babysitting Service

Frank Furster, a 36-year-old small-business owner and his 17-year-old wife, Iliana, operated the Country Walk Babysitting Service out of their Miami home. Parents became concerned because of numerous problems with their children. One parent believed her child had been drugged and abused by the Fursters; other children claimed that Frank and Iliana kissed their penises, inserted fingers into their rectums, and paraded nude in front of them. Interweaved among the credible allegations that the children made were ones that seemed fabulous, such as riding on sharks and eating the head of another person. The children claimed that Frank Furster videotaped their sexual abuse, although the alleged tapes were never found. In 1986, 3 years after parents first voiced their concerns, Frank and Iliana Furster were tried on multiple counts of child abuse, rape (sexual battery), sodomy, terrorism, and lewdness with a child. The children told interviewers about events that allegedly had taken place several years earlier, when they were aged 1–5 years. After nearly 15 months of denials, Iliana Furster turned state's evidence against her husband. She revealed that she too had been a victim of Frank's abuse and corroborated many of the claims the children made. Frank Furster was convicted of 14 counts of sodomy, rape, and abuse and was sentenced to the equivalent of several life sentences. Iliana was sentenced to 10 years, with 10 additional years of probation.

From Case Studies to Systematic Research

These two cases highlight different aspects of children's credibility that have been the focus of research. The first aspect involves the accuracy of recalling events over long periods of time. In the Country Walk case, the children sometimes described events that allegedly occurred several years before they gave their testimony. Hundreds of studies have examined the degree to which children are able to accurately encode, store, and retrieve different types of information. Most of these studies, however, have examined short-term recollections of objects (as opposed to actions) and of peripheral (as opposed to central) events. Despite these limitations, on the basis of this literature it is safe to conclude that memory skills do improve with age (e.g., see reviews by Kail, 1989; Ornstein, 1978; Schneider & Pressley, 1989).

Notwithstanding this age-related improvement in recall, even very young children's memory is accurate over long delays if the materials and procedures make sense to them (Flavell, 1985) or if the object to be remembered is a salient action or a personally meaningful event (Cutts & Ceci, 1988; Fivush & Hamond, 1990; Jones, Swift, & Johnson, 1988; Perris, Myers, & Clifton, 1990). Recall of action-related events is highly reliable, even in preschoolers (e.g., Davies, Tarrant, & Flin, 1989; Jones et al., 1988), particularly when they are participants in an event (Rudy & Goodman, 1991).

Furthermore, age differences in *recognition* memory are far less pronounced than age differences in *free recall*, and at times these are nonexistent (Ceci, Ross, & Toglia, 1987; Cole & Loftus, 1987; Jones et al., 1988; List, 1986; Saywitz, 1987). For example, preschoolers remember as much as adults when the task does not emphasize verbal recall (Nurcombe, 1986) and in response to specific questions. Even 3-year-olds recognize as many familiar drawings as 12-year-olds (Ceci et al., 1987). Studies such as these indicate that preschoolers' recognition memory can be remarkably accurate (Kail, 1989).

The second aspect of children's testimonial credibility concerns their "suggestibility," and it is this aspect that is the focus of our review. According to its broadest definition, *suggestibility* concerns the degree to which children's encoding, storage, retrieval, and reporting of events can be influenced by a range of social and psychological factors. This broad definition contrasts with the narrower and more traditional definition of suggestibility, which asserts that it *is* "the extent to which individuals come to accept and subsequently incorporate post-event

information into their memory recollections" (Gudjonsson, 1986, p. 195; see also Powers, Andriks, & Loftus, 1979). This narrower definition implies that suggestibility can only be unconscious (i.e., interfering information is unwittingly incorporated into memory); suggestibility results from the provision of information following an event as opposed to preceding it; and suggestibility is a memory-based, as opposed to a social, phenomenon. We adopt the broader definition of suggestibility because it implies that (a) it is possible to accept information and yet be fully aware of its divergence from some originally perceived event, as in the case of "confabulation" (such as is shown by brain-injured patients; see Johnson, 1991), acquiescence to social demands, or lying (see footnote 5); thus, these forms of suggestibility do not involve the alteration of memory. (b) Suggestibility can result from the provision of information preceding *or* following an event. (c) Suggestibility can result from social as well as cognitive factors. Thus, this broader view of suggestibility is consistent with the legal use of this term to connote how easily one is influenced by both subtle suggestions and leading questions, as well as by explicit bribes, threats, and other forms of inducement.

Within this framework, one can examine how much children's testimonies reflect their desire to protect themselves, the cultural and personal beliefs that also influence adults' willingness to accept children's testimony, and the nature of the interrogations that induce children to make certain statements or accusations. For example, in our two sample cases, expert witnesses and prosecutors insisted that the children must be believed because children do not lie and they cannot be mistaken about sexualized claims. In the Country Walk case, there was repeated provision of an atmosphere of accusation, with interviewers informing children, "It's okay to tell. . . . You'll feel better once you tell." Finally, in the Wee Care Nursery School case, most of the children were told by interviewers prior to their own disclosures that their peers had already disclosed that Kelly Michaels was a bad person who had hurt them. These are issues that we return to in evaluating the research on children's incorporation of adult beliefs and the creation of an "atmosphere of accusation" in interviews.

By broadening the definition of suggestibility to entail nonmnemonic influences, we summarize the literature on the following two questions: First, are younger children more suggestible than older children? Second, to what degree does suggestibility reflect cognitive, social, and biological factors? The examination of these questions allows for a more precise understanding not only of the conditions under which children are suggestible but more generally of the causal mechanisms that underlie their suggestibility.

Before turning to these issues, it is important to emphasize that we do not mean to imply that adults are not suggestible, that their memories are always reliable, or that their testimonies are highly accurate. These statements are clearly false. There is a sizable literature on the suggestibility of adults' memory (e.g., Belli, 1989; Gudjonsson, 1986; Lindsay, 1990; Loftus, 1979). In this article we examine factors that may influence witnesses of all ages but that may exert a disproportionate influence on children.

Historical Review

Early Research: 1900–1914

Historically, interest in children's testimonial competence, both by the legal profession and by social scientists, has reflected specific judicial events, the structure of the judicial system, and general social conditions of the era. In the United States, there was little interest in this field until the last half of the 20th century. To some degree this reflected the Salem Witch Trials of 1692. At that time a group of children gave false testimony in the witchcraft trials of more than 20 residents of Salem Village and Salem Farms. The girls made fantastic claims (Ceci, Toglia, & Ross, 1990). Several years after the execution of defendants, some of the child witnesses publicly recanted their testimonies. For the most part, the prevailing legal attitude for the following 300 years has been one of skepticism about the testimony of child witnesses (e.g., Wigmore, 1935). Repeatedly, legal scholars have cited the excesses of Salem as a basis for their views of child witnesses.

Although there was little if any interest among psychologists in children's testimonial competence in the United States at the start of the 20th century, this was not the case in Europe, where systematic research on adults' and children's testimony flourished, especially in the Federal Republic of Germany and France. To a large degree, differences in the adjudication procedures in the two continents can account for these differences in research in this area. An inquisitorial system of justice prevails in many European countries in which the judge is responsible for calling and questioning witnesses. Because there is often no jury, the European judge is more likely to call on expert witnesses to testify about the competence of witnesses (Loh, 1981). In the early part of this century, these expert witnesses were often psychologists who carried out experiments to examine the veracity of the children's testimony. By contrast, in an adversarial system, such as the one used in the United States, the use of opposing attorneys and a jury was considered sufficient to evaluate witness credibility (see Loftus, 1986, for additional details).

Because few of the early studies on testimonial competence were published in English, unilingual Anglophones have had to rely on reviews of this research for its details. The most influential of these were published in the *Psychological Bulletin* by Whipple (1909, 1911, 1912, 1913). These reviews were notable for their coverage of the child suggestibility research by European psychologists and medical experts. In the course of these reviews, Whipple became increasingly convinced that young children are highly suggestible and capable of making serious errors in their testimony, even when they testify about matters of great personal importance. Although these reviews are still cited as definitive summaries of early research (e.g., Baxter, 1990; Goodman, 1984a), they provide few details of the actual procedures or results of the studies. This is unfortunate for today's reader because some of the methodologies used in modern research were developed by these early scientists who also had sophisticated views on issues that are currently debated. In order to introduce the reader to some of these methodologies and issues, we provide some details on the work of four pioneer-

ing European scientists: Binet, Stern, Varendonck, and Lipmann. The following summaries are based on our translations or published translations of primary source materials.

A. Binet. On the basis of a series of studies of children between the ages of 7 and 14 years, Binet (1900) claimed that suggestibility reflected the operation of two classes of factors. The first class concerns the influence of a prominent thought (autosuggestion) that develops within the individual, and is not the result of another's influence, but that paralyzes the critical senses. The second class of factors is external to the individual and reflects mental obedience to other individuals.

Although Binet's (1900) autosuggestion techniques were adapted by the next generation of researchers, they are rarely used in modern studies. The best known of these involved showing children a series of lines and then asking them to draw the final one. The first five lines progressively increased in length. The sixth line, however, was the same length as the fifth. Children tended to be swayed by the perceptual or internal suggestion of ever-increasing lines; thus, their drawings of the sixth line were too long. However, the influence of the suggestion was not long-lasting; children could easily regain control of themselves and accurately redraw the target line when asked to do so at the end of the experiment.

In contrast to Binet's (1900) paradigms for examining internal forces, his paradigms to examine external forces are still used today. In one study, children saw five objects for 10 s (e.g., a button glued to poster board). Some were told to write down everything they saw. Others were asked direct questions (e.g., "How was the button attached to the board?"). Others were asked mildly leading questions (e.g., "Wasn't the button attached by a thread?"). Some were asked highly misleading questions (e.g., "What was the color of the thread that attached the button to the board?"). The major finding was that free recall resulted in the most accurate statements and that highly misleading questions resulted in the most inaccurate statements. As we discuss later, this pattern of results has since been replicated in dozens of studies. Children's answers to Binet's questions were characterized by an exactness and certainty, regardless of their accuracy. Because children did not correct their inaccurate responses, Binet concluded that their erroneous responses reflected gaps in their memories; they attempted to fill in these gaps to please the experimenter. However, once an erroneous response was given, Binet proposed that it became incorporated into memory. In other experimental contexts, Binet directly ascribed children's suggestibility to social factors, namely, children's eagerness to comply with adult suggestions rather than to memorial factors. In those cases, Binet discovered that children's suggestibility was not long-lasting; they quickly realized their errors.

In sum, Binet was prescient in three ways: First, he distinguished between errors of reporting caused by actual memory changes versus those caused by social conformity, arguing that the latter include attempts to please adult authority figures and do not always reflect incorporation of the suggestion into the memory record. Later, we review modern evidence on the debate over the supremacy of cognitive versus social mechanisms. Second, Binet foreshadowed the current debate over whether the original memory trace is itself impaired or simply allowed to "coexist" with traces of the erroneous suggestion (Loftus, 1979). Third, Binet alerted researchers to the weak relation between confidence and accuracy (see Bothwell, Deffenbacher, & Brigham, 1987, for current data).

W. Stern. Stern (1910) developed two types of experiments that are still in use today. In the first paradigm, subjects were shown a picture and asked to study it for a short period of time. Immediately after its presentation, they were asked to recall what they had seen. They were then asked a series of direct questions, requesting information that was in the picture, and a series of misleading questions, requesting information about nonexistent objects. In one study that included 7- to 18-year-olds, free recall produced the fewest errors, whereas misleading questions produced the most errors (Stern, 1910). Although younger children were the most suggestible, even the 18-year-olds occasionally were misled by the suggestive questions.

The second paradigm, the "reality" experiment, was developed to mimic situations that were closer to real life. Here, naive subjects observed staged incidents. In a typical experiment, an argument occurred during a seminar between two students, one of whom drew a revolver. The other students in the class were then questioned about the scenario.

Stern made several observations that continue to be important. He warned about repeated questioning of the same event, claiming that a subject's original verbal answers are better remembered than the actual events themselves (Stern, 1910). He also talked about the "force" that questions have in determining answers, claiming that many children answer questions because they view them as imperatives. Stern argued that the questioner, by virtue of the nature of the questions asked, is often responsible for the unreliable testimony of witnesses. Finally, Stern (1910) believed that children are especially suggestible at certain times of their lives when they merge fiction and reality. Children, particularly girls, were said to be suggestible around puberty as the result of hormonal changes. Stern is to be credited for illuminating the notion of "reality monitoring judgments," an area of continued activity (Johnson, 1991; Johnson & Raye, 1981), although his predictions concerning both age and sex effects were subsequently shown to be wrong.

J. Varendonck. Varendonck, a Belgian psychologist, was an expert witness in a trial involving allegations by several children that a young girl named Cecile was murdered by a local man (Varendonck, 1911). Two of Cecile's friends who had played with her on the day of her murder were awakened that night by Cecile's mother to ask of her whereabouts. One of the children replied that she did not know. Later that night, she led the police to the spot where the children had played, not far from where Cecile's body was found. In the next month, the two children were repeatedly interviewed by authorities who asked many suggestive questions. The children quickly changed their original testimony of not knowing about Cecile's actions on the day of her murder. They provided details of the appearance of the murderer as well as his name. Because of an anonymous letter, the police arrested the father of one of the playmates for the murder of Cecile. On the basis of the details of the case, Varendonck was convinced of the defendant's innocence. He quickly conducted a series of studies with the specific intent of demonstrating the unreliability of children's testimony.

In one study, Varendonck (1911) asked the children in his class to describe a person who had approached him in the school yard that morning. Although there was no such person, most of the children fell sway to his suggestion, with 17 out of 22 giving a name for the person, the color of his clothes, and so on. Varendonck claimed that the types of questions he used were parallel to those that the examining magistrate used with one of the child witnesses.

Varendonck concluded from his demonstrations that the two children's statements to the police were false, the result of suggestions provided by influential adults. He carefully documented how the children changed their testimonies between the first and second interrogations and how other social factors conspired to produce their testimony. He concluded that children cannot observe accurately and that their suggestibility is inexhaustible ("We cannot set the least value in their declarations"; Varendonck, 1911, p. 168). His work is noteworthy because of the direct forensic applications of his empirical data.

O. Lipmann. The work of Lipmann, a German psychologist, is of interest because many of his hypotheses are the focus of modern research. Consistent with Binet, he concluded that cognitive as well as social factors accounted for children's greater suggestibility. Children were thought to have different, not fewer, memories than adults because they were sensitive to different attributes of stimuli than adults. When children are questioned by adults, who have authority over them, about events that are neither essential nor salient to the child, the child will attempt to revise his or her memory, making the report consistent with the question. "If the respected person who is questioning me expects such an answer then it must be the right one" (Lipmann, 1911, p. 253). Thus, rather than answering "I do not know," the child accepts any material that comes to mind to fill in these gaps, whether it is imaginary or real. Eventually everything that is imagined becomes real (i.e., the child fails to differentiate fantasy from reality). Modern researchers would return to the issue of the young child's ability to separate the sources of their information, including whether it was imagined or perceived (Foley & Johnson, 1985; Foley, Johnson, & Raye, 1983; Lindsay, Johnson, & Kwon, 1991). Modern researchers would also return to the idea that children have different perceptions (or scripts) of the world than adults and that these can also affect the nature of their memories.

Summary of research during the early European period. Two important elements of the early European work on children's suggestibility deserve mention. First, all of the researchers during this early period were interested in applications of children's memory research to the legal system. Second, multifactorial mechanisms underlying suggestibility were posited. These involved cognitive factors related to children's encoding, storage, and retrieval of events as well as social factors related to children's compliance with authority figures. It should also be noted that this early work foreshadowed a large number of findings that were to appear in the modern literature, such as the idea that repeated questioning is detrimental, that questions are interpreted as "imperatives" by young children, that free recall produces fewer errors than yes–no questioning, that fantasy–reality distinctions are problematic for very young children, and that even adults are suggestible.

The Dry Middle Years: 1924–1963

Overview. Although European courts were eager consumers of the psychological research on children's suggestibility, the same could not be said of American courts. According to Loh (1981), similar studies of the reliability of witnesses in the United States were rejected by the legal profession. Münsternberg (1907a, 1907b), a Harvard psychologist, summarized the European literature on the unreliability of adult witnesses and made a strong case for using psychological methods in U.S. courts of law. His position, however, was ruthlessly criticized by jurists such as Moore (1907, 1908) and Wigmore (1909) on the grounds that psychology had nothing useful to offer law. Wigmore claimed that psychological experimentation produced results based on group averages, whereas in a court of law the relevant issue concerns the reliability of a specific witness in a specific situation. (Wigmore did soften his stance against psychology later in his career.) This rejection of psychological research by leading members of the U.S. legal community resulted in a long hiatus, during which little work was carried out by psychologists on the accuracy of witnesses' testimony. Until the reemergence of this genre of research in the late 1970s, there was only a handful of studies on children's suggestibility, most carried out in the 1920s and 1930s and, for the most part, marked by their unoriginality. The major focus of these studies was to examine the relations of age, intelligence, and sex to suggestibility or to examine the correlations among different suggestibility measures, most of which were adaptations of tasks devised by Binet and Stern. The interesting questions raised by Binet, Stern, Varendonck, and Lipmann went unaddressed. For this reason, we do not devote as much space to these studies.

Otis (1924) examined the development of children's ability to rely on their own judgments. Her test included many items similar to those devised by Binet to assess autosuggestion. Other questions, which were phrased in a leading manner, assessed the influence of external forces. Students in Grade 3 through college were tested. Suggestibility decreased as a function of age and intelligence. Using a similar measure, Hurlock (1930) replicated these results with a sample of 10- to 17-year-olds. Burtt and Gaskill (1932) showed students in Grade 4 and college a movie and asked them leading and nonleading questions about what they had witnessed. College students' errors on the suggestive questions were much lower than those of the fourth graders.

Sherman (1925) examined the association of suggestibility with chronological and mental age in normal and mentally challenged children. The children were given eight different tasks that involved "direct" questions and "auto-suggestions." In general, suggestibility decreased with age in both samples. In addition, suggestibility also decreased as mental age increased in the mentally challenged sample. By contrast, Messerschmidt (1933) tested 6- to 16-year-olds on a battery of similar tests and found a strong association between age and performance that was consistent across tasks. The youngest children were the most suggestible across tasks, and the oldest children were the least suggestible.

The next suggestibility study involving children did not ap-

pear in the literature until 30 years later (McConnell, 1963). Several measures of visual perceptual suggestibility were given to children in Grades 1–12. For example, they were shown two equal objects and asked to circle the one that looked the biggest. Next, the experimenter told them that one of the circles really was larger than the other and to circle the largest. They were given the option of marking "neither." On all tasks suggestibility correlated with age.

Summary of research during the dry middle years (1924–1963). Two consistent findings emerge from this set of studies. First, younger children were more suggestible than older children and adults. Second, there was a negative correlation between suggestibility and IQ, with those possessing lower IQs being less able to resist suggestion. However, it must be kept in mind that in most cases, many of these memory measures were paper-and-pencil tests; thus, the correlations with IQ may reflect the fact that the poorer students had more difficulty dealing with written materials, or keeping their attention focused during long written tasks, rather than with suggestibility of the experimental manipulations.

In contrast to the earlier European studies, the studies conducted by American researchers during the 1924–1964 period were not couched in legally relevant terms. For reasons stated earlier, there was never any mention of the applicability of these findings to the issue of children's testimony. One is also struck by their atheoretical nature. No new principles or paradigms were discovered, and there was little theorizing about the underlying causes of developmental differences in suggestibility. One issue that does recur is the degree to which suggestibility is a trait. However, not only are the data inconsistent across studies, but even on those occasions when the same patterns of data were reported, they led to different interpretations. One view was that susceptibility to suggestive questioning resulted from a traitlike tendency (e.g., Aveling & Hargreaves, 1921). Children were more suggestible because of immature but developing mechanisms that made them more susceptible to external factors. According to this view, suggestibility was an individual-differences variable along which people could be differentiated and along which children as a group were relatively deficient. Others (e.g., Remmers, Cutler, & Jones, 1940; Sherman, 1925) viewed suggestibility not as a trait but as a function of task-specific factors, including characteristics of the experimenter and laboratory.

Our review suggests that a consensus was building about children's testimonial incompetence, reflected in Burtt's (1948) description of children as "dangerously vulnerable to coaching and erroneous leading questions: Suggestion is especially apt to play a role in the testimony of children because they are more suggestible than adults" (p. 307).

The Modern Period: 1979–1992

Following a 16-year hiatus in research on children's suggestibility, the late 1970s marked a resurgence of interest among developmental researchers in the reliability of children's reports. Since 1979, more than 100 studies on children's suggestibility have been reported. Four interrelated factors account for this dramatic increase in empirical work.

First, there has been a broadening of admissibility of expert psychological testimony in recent years, particularly with regard to issues concerning mental disorders, pretrial publicity, and civil rights (see Loh, 1981). Thus, social science research, after a long period of being ignored or rejected by judicial policymakers, has come to be viewed, at least on occasion, as being relevant to the legal system. Second, in part fueled by the sociopolitical *zeitgeist* of the late 1960s, social scientists attempted to apply their scientific training to socially relevant issues, particularly those concerning children's rights and the protection of minors. Third, many studies were motivated by or influenced by methods and theories emanating from studies on eyewitness testimony of adults, which, for the reasons just mentioned, were also increasing in number.

The fourth and undoubtedly the biggest stimulus for the explosion of research on children's suggestibility is the legal community's heightened interest in behavioral science data related to specific innovations for dealing with child witnesses. For example, until recently, there has been a reluctance to accept the uncorroborated statements of child witnesses in courts of law in all English-speaking countries (Chadbourn, 1978). This reluctance is reflected in competency hearings, corroboration requirements, and cautionary instructions that some North American judges give to juries concerning the inherent reliability risks of convictions based solely on the testimony of child witnesses (Andrews, 1964; Cohen, 1975). However, since the 1980s, more children are being admitted as witnesses as a result of dramatic increases in reports of crimes involving sexual abuse and physical abuse in which the child has been a victim or a witness. In 1989, there were 2.4 million reports of suspected child maltreatment in the United States; 900,000 were substantiated (Daro & Mitchel, 1990).

As a result of the ineffective prosecution of child abuse cases, in the past decade the legal system has been forced to change some of its rules concerning the admissibility of child witnesses' testimony. During the 1980s all states dropped their corroboration requirement for children involved in sex abuse cases, a crime that by its nature is often without 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 witness. As more and more children are allowed to provide uncorroborated testimony, courts begin turning to psychological research to inform their proceedings.

Because children are increasingly being admitted as courtroom witnesses, courtroom procedures have also been modified. Of particular pertinence to this article, most states have evidentiary codes that permit asking the child leading questions in sex abuse cases. Other procedures, such as shield laws and hearsay exclusions, have been instituted to assist child witnesses (see McGough, in press). In light of claims that such modifications challenge the constitutional rights of defendants (*Maryland v. Craig*, 1990), it is important to obtain empirical data that such procedures do in fact enhance the court's truth-seeking function.

This increased demand for scientific data on children's credibility has resulted in a large number of recent studies that are methodologically superior to the older work and that aim for

greater external validity through the use of experimental procedures that seem more realistic. Thus, in contrast to many of the older studies that required children to make perceptual judgments (e.g., Which line is longer?) or to recall neutral stories or pictures, many of the newer studies have examined the manner in which children process and recall important, personally experienced, highly salient, affectively loaded events in the context of strong preevent or postevent suggestions. However, although much of this research on children's recollections is being carried out in more naturalistic contexts, this does not in itself make it generalizable to a particular court case unless the research context closely mirrors the factors "at bar" (see Ceci, 1991; Loftus & Ceci, 1991).

The current research is also beginning to reexamine (and in some cases reinvent) hypotheses that were first raised by the early European scientists. The focus has thus shifted from simply examining whether children are suggestible to determining under what conditions they are suggestible. To some degree this shift has been influenced by current work on the testimonial competence of adults (e.g., Melton & Thomson, 1987), as well as by recent basic research on the cognitive and social development of children.

A final feature of the newer studies concerns the ages of the children studied. In contrast to previous studies, which focused on school-aged children, modern researchers often include preschoolers. Because preschoolers are increasingly being called to testify, the need for a greater understanding of their testimonial accuracy is urgently needed. Approximately 25 of the studies described in this article involved preschoolers; by contrast, during the first 80 years of this century, there was not a single study, to our knowledge, that included children this young.

We turn next to a review of the modern child suggestibility literature. During the modern period, some investigators, like their predecessors, have emphasized evidence of children's special vulnerability to suggestions. Other investigators, however, have emphasized evidence of children's ability to resist suggestions and to give accurate testimony. We describe five representative studies from each of these positions. We selected these studies because each has been cited in support of the claim that there are or are not developmental changes in children's suggestibility. As we show, studies published by both camps often contain mixed results (i.e., there is evidence of age-related changes under some conditions but not under others). Furthermore, there are inconsistencies in the pattern of results between some studies. Such inconsistencies illustrate our earlier claim that despite the superior methodology and greater ecological realism, these modern studies have initiated and fueled, rather than resolved, disagreements among researchers over the suggestibility of children's statements.

Children are more suggestible than adults. Our review of the studies conducted during the first 70 years of this century indicated that almost without exception, researchers believed children were more suggestible than adults. The following five examples of recent research, which used more sophisticated methodologies and ecologically realistic settings, showed results similar to the earlier work.

1. Cohen and Harnick (1980) presented a 12-min film about a petty theft to 9-year-olds, 12-year-olds, and college students and tested their memory for the details of the film immediately afterward and 1 week later. For the first interview, half of the 22 probe questions were misleading (e.g., "The young woman was carrying a newspaper when she entered the bus, wasn't she?"), and the other 11 questions were not phrased in a misleading manner. The youngest subjects produced the least accurate responses to both nonmisleading and misleading questions, indicating that they were more suggestible than the older subjects. These age differences were not reliable, however, when the subjects were tested 1 week later, using a multiple-choice question format. The authors concluded that younger children were more likely to consciously submit to suggestions than older subjects but that the suggestions did not differentially affect their memory for the event.

2. King and Yuille (1987) staged a live event for 6-, 9-, 11-, and 16-year-olds. The subjects were seated in a room when a stranger entered to care for some plants. Prior to leaving the room, the stranger noted the time and indicated it was late. When the children were subsequently interviewed, they were asked for a description of what they could recall as well as some leading questions such as "On which arm did the man wear his watch?" (He had not worn a watch.) The 6-year-olds were significantly more suggestible than 9- to 16-year-olds, and they also recalled less.

3. Ceci et al. (1987, Experiment 1) presented short stories accompanied by illustrations to 3- to 12-year-olds. One day after the presentation, they provided misleading information about aspects of the stories to half of the subjects. Two days later, they tested the children's memories of the stories by having them select from a series of four pictures the two that had actually appeared in the story. Age differences were obtained only for children who were given misleading information. Preschoolers were more likely than the older children to select pictures that were described in the misleading session than pictures that appeared in the actual story.

4. Ornstein, Gordon, and Larus (1992) tested 3- and 6-year-olds' memories of a pediatric examination. Approximately half of the children at each age were tested immediately following the examination and 1 week later, and the others were tested immediately and 3 weeks later. Most of the children were asked some misleading questions. At each test session, the older children's memories were better than the younger children's as assessed by free-recall and objective questions. Furthermore, the 3-year-olds gave fewer correct answers to the misleading questions than did the 6-year-olds during the first two testing periods. These age differences were not reliable after 3 weeks; this reflected the fact that 6-year-olds' accuracy on misleading questions was greatly reduced between the first and last session relative to that of the 3-year-olds.

5. Oates and Shrimpton (1991) studied the effect of questioning on the memories of two groups of 4- to 12-year-olds. One group received a blood test and the second group encountered a friendly stranger in their school library who put a loose cotton shirt over the child's clothes and then removed it. The children's memory of the event was assessed 4–10 days following the event or 3–6 weeks later. On all measures, children in the blood group performed comparably to children in the library group. Also, children were more accurate when tested after the short

delay than after the long delay. Of particular importance, older children (aged 7–12 years) performed better than younger children (aged 4–6 years) on free-recall, direct questions, and some types of misleading questions. Compared with younger children, older children were less misled about actions, but there were no age differences on resistance to being misled about the person with whom they interacted. Finally, the effect of the delay of interview was especially consequential for the misleading *action* questions; children interviewed after a long delay were more susceptible to suggestion than those interviewed after a short delay. Recently, Poole (in press) has found that long delays (nearly 2 years) are disproportionately more detrimental to the memories of 4-year-olds than older children and adults.

Younger children are not less suggestible than older children. In view of the findings presented to date, it is surprising to discover that there are those who argue that there is no evidence of age differences in suggestibility. For example, three years ago Gary Melton, the past president of the American Psychological Association's Division on Psychology and the Law, expressed concern over the fact that the dissenting Supreme Court justices in the case of the *State of Maryland v. Sandra Ann Craig* maintained that children were substantially more suggestible than adults. Melton asserted that "the dissent's discriminate plucking of such material from the psychological literature doesn't reflect the broad findings within the field" (cited in DeAngelis, 1990, p. 1). More recently, Melton (1992) reaffirmed (and seemingly strengthened) this assertion:

> 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)

Melton's dismay reflects the fact that there are studies to support the view that children are no more suggestible than adults. The following five studies are examples of this literature.

1. Marin, Holmes, Guth, and Kovac (1979) exposed 5-, 8-, and 12-year-olds and college students to a live staged argument between two adults. After a brief delay, subjects were asked 20 objective questions and an additional misleading question. The impact of the misleading question was assessed 2 weeks later, when all 21 questions were asked in a nonleading form. Children did not differ from college students on objective questions asked immediately after the event. Furthermore, although the introduction of the misleading question produced a significant increase in inaccurate answers on the corresponding objective question asked 2 weeks later, the size of this suggestibility effect was similar across all ages. Thus, children were no more suggestible than adults.

2. Duncan, Whitney, and Kunen (1982, Experiment 2) showed 7-, 9-, and 11-year-olds and college students slides depicting scenes from the movie *Star Wars*. Following the presentation of the slides, subjects received related, unrelated, and neutral information that was either consistent with the slides they had just seen or was consistent with distractor slides that were shown at the time of testing. In a complex analysis that entailed having subjects' d' recognition scores contingent on their memory criterion (i.e., using only the stories on which children demonstrated good memory for follow-up questions), they showed that the older subjects were *more* likely than younger subjects to incorporate misleading verbal information into their visual memories for slides.

3. Flin, Boon, Knox, and Bull (1992) exposed 6-year-olds, 10-year-olds, and adults to a realistically staged argument during a presentation on foot hygiene by a nurse in the school auditorium. Half of the subjects were questioned about the event 1 day later, and all subjects were questioned 5 months later. Three of the questions contained erroneous suggestions. In both interviews, responses to these questions were highly accurate across all age groups; few subjects of any age accepted the erroneous information.

4. Perhaps no researcher has done more to redress the historical imbalance in favor of child witnesses than Gail Goodman. After almost a century of research criticizing and belittling the accuracy and suggestibility of child witnesses, Goodman has presented a far more optimistic picture of children's abilities. Her work is animated in part by a desire to know whether nonabused children will make false claims of abuse in response to erroneous suggestions by adults. In order to examine this question, her strategy has been to interview nonabused children about sexual as well as nonsexual experiences.

As one example, Rudy and Goodman (1991) studied pairs of 4- and 7-year-olds who were left in a trailer with a strange adult. One child played a game with the adult that involved being dressed in a clown's costume and being lifted and photographed while the other child was encouraged to carefully observe this interchange. Approximately 10 days later, the children were asked suggestive and nonsuggestive questions about the event. Some of these questions concerned actions that might lead to an accusation of child abuse, such as "He took your clothes off, didn't he?" Across all question types, there were few differences between participants' and bystanders' responses.[2] Older children were more accurate than younger chil-

[2] Note that this statement differs from the conclusions offered by Rudy and Goodman: "As predicted, participation in a real-life event heightened the children's resistance to suggestion. On misleading questions, participants were less suggestible than bystanders. On misleading questions concerning the confederate's appearance, 4-year-old participants were less suggestible than 4-year-old bystanders, and an age difference appeared only for bystander witnesses. This pattern indicates that participation can strengthen resistance to suggestion and that at least at times, the effects are especially evident for young children" (Rudy & Goodman, 1991, p. 534). Rudy and Goodman failed to consider in this discussion that when "don't know" answers were included in the data, only one of the four analyses of misleading questions yielded significant results for participation. When only the misleading abuse questions were considered (a fifth analysis), there was no significant effect for participation. Furthermore, the analysis of the nonmisleading direct questions and of the free-recall data failed to reveal any advantage for participation. Thus, their conclusions concerning the effects of participation seem overgenerous, given the actual pattern of results.

dren on nonsuggestive (abuse and non-abuse-related) questions. On misleading questions, these same age effects were obtained only for the nonabuse questions. Accuracy rates on the abuse misleading questions were similar for the younger and older children. A more detailed analysis of the incorrect answers to the suggestive abuse questions revealed only one false report of abuse; a 4-year-old bystander falsely claimed that he and the participant had been spanked.

5. A second study conducted by Goodman and her colleagues (Goodman & Clarke-Stewart, 1991; Saywitz, Goodman, Nicholas, & Moan, 1991) examined 5- and 7-year-old girls' memories of medical examinations. Half of each age group had a scoliosis exam, and half had a genital exam. Children were tested 1–4 weeks following their exam. Children were asked suggestive and nonsuggestive questions that were abuse related or non-abuse-related. The older children's answers to the suggestive non-abuse questions and to the nonsuggestive abuse questions were more accurate than those of the younger children. However, there was essentially no difference in resistance to suggestibility for suggestive abuse questions (e.g., "How many times did the doctor kiss you?"), with few children giving incorrect responses. The 7-year-old children never made a false report of abuse, and this occurred only 3 out of 215 times for the 5-year-olds.

Weighing the pros and cons. Because there was so much variability in the methodologies used in the 10 studies, it was not possible to resolve the modern controversy concerning age differences in suggestibility by direct comparisons of them. That is, no two studies were alike on many of the relevant dimensions, such as the nature of the event to be remembered (e.g., verbal stories, slide shows, or physical examinations); the timing of the misleading information (prior to the memory test vs. during it); timing of the interview (shortly after the stimulus event vs. several weeks after it); and the type of data-analytic techniques (analysis of raw data vs. signal-detection techniques).

It also does not seem fruitful to resolve the existing controversy through a point-by-point criticism of the methodological weaknesses of each study. Methodological concerns can be raised with equal force at studies on both sides of the debate. We now provide a sampling of some the concerns that, although not exhaustive, demonstrates that interpretative problems plague studies on both sides of the debate. For example, the failure to find statistically reliable age differences on a number of the suggestibility measures may reflect the use of relatively small sample sizes, which masks real developmental differences. As an example, Cohen and Harnick's (1980) failure to find reliable age effects after a 2-week delay could have been caused by their having only 3 subjects in each cell of their analysis. This could have prevented observed age differences that were large in magnitude from reaching traditional levels of reliability. (Incorrect response rates to misleading questions were 51%, 33%, and 22%, respectively, for the three age groups.) Examination of the sample sizes and large variances reported in the Duncan et al. (1982) and Flin et al. (1992) raise similar concerns.

A second concern is the number of suggestive questions included in the interviews. For example, the Marin et al. (1979) study included only one leading question. Because chance accuracy with a single yes–no question is .50, it is noteworthy that the rate of answering the suggestive question correctly after a 2-week delay was essentially this value for the four age groups (.50, .50, .46, and .46, respectively). These data leave open the possibility that all subjects might have been influenced by the misleading question but that floor effects prevented a powerful test of any age difference. Concerns about the number of suggestive questions can also be raised for studies that showed age effects. In the Ornstein et al. (1992) study, the size of the question set changed for each child. It is possible that older children were less suggestible because they were asked more suggestive questions (producing a larger denominator and a smaller overall suggestibility ratio).

Perhaps age differences are obtained only when situations are highly artificial or irrelevant to forensically important issues. For example, in the Ceci et al. (1987) study, the experimental context was a nursery school story in which unfamiliar characters were described to children for brief periods and later described erroneously. That young children succumbed to such suggestions under those circumstances does not necessarily indicate that they will do so in response to more emotionally salient and powerful materials. Nevertheless, age trends in suggestibility effects have been reported for more stressful and naturalistic situations (e.g., Oates & Shrimpton, 1991; Ornstein et al., 1992). And, in their ecologically based studies of thefts, both Cassel and Bjorklund (1992) and Warren and Hagood (in press) found age differences in succumbing to suggestive questions even for central events, with younger children more suggestible.

The linguistic complexity of the misleading questions may be related to the appearance or nonappearance of age-related differences in suggestibility. Some of the questions used in various studies might have been too complex and beyond the comprehension of young children. An example of such a question is as follows: "What did the costume that he asked the other boy to wear look like?" (Rudy & Goodman, 1991, p. 538). In response to such questions, the children might have answered "I don't know" (which was counted as an accurate answer), not because they were resisting the suggestion but because they did not comprehend the question. This could obviate potential age-related differences, particularly if the "don't know" answers of the younger children reflect poor comprehension, whereas the "don't know" answers of the older subjects reflect resistance to suggestion. However, similarly difficult questions were also found in studies that did report age-related differences in suggestibility (e.g., from Oates & Shrimpton, 1991, p. 8): "The person who gave you the blood test put your arms behind your back, didn't she?"), forcing the alternative argument that perhaps high rates of acquiescence reflect poor comprehension of the questions.

One might also note that although each of the 10 studies cited are commonly used to provide evidence for or against age differences in children's suggestibility, within each study there are conflicting results. Thus, Rudy and Goodman (1991) consistently reported that there were age-related differences in children's answers to misleading questions, except for one special type of question. Similarly, Cohen and Harnick's (1980) study is

commonly used to support the age-difference position, but age-related differences were obtained only on the first, and not on the second, testing.

This discussion demonstrates that any attempt to resolve these inconsistent results by a point-by-point examination of whether researchers on the two sides of the debate use different age groups, settings, and events still leaves many contradictory findings and does little to illuminate the nature of the age differences when they do occur. Thus, rather than attempting to contrast each of these studies on a microlevel, it seems more fruitful to explore the causal mechanisms that may underlie obtained suggestibility effects and, in turn, to consider how these various mechanisms might explain the appearance or nonoccurrence of age trends in suggestibility effects. With this goal in mind, we now explore three types of factors: cognitive, social–motivational, and biological.

Causal Mechanisms: Cognitive Factors

Children become increasingly cognitively sophisticated with development as a result of a confluence of attainments in memory, concept formation, reasoning, language ability, and introspective awareness of the cognitive system's executive functions (Ceci, 1990). In this section, we discuss the aspects of this research that have the greatest relevance to understanding the potential causal mechanisms of suggestibility.

Memory

One issue related to the underlying mechanisms of suggestibility involves the extent to which erroneous postevent information interferes with the original memory. We describe the procedures developed by Loftus and her colleagues (e.g., Loftus, Miller, & Burns, 1978) to examine these effects first, because the developmental data on suggestibility are primarily based on these procedures or on modifications of them. Subjects first view an event that consists of a number of details (e.g., a man holding a hammer while drinking cola). They then receive information about the event, some of which is misleading (e.g., the man was holding a wrench while drinking cola). Finally, their memories for the original events are tested (e.g., Was the man holding a hammer or a wrench?). Commonly, subjects make more errors for items about which they were given incorrect information than for control items (e.g., Ceci et al., 1987; Loftus et al., 1978; Marin et al., 1979). Thus, although they incorrectly reported that the man was holding a wrench, they correctly remembered that he was drinking cola.

Although this demonstration of the phenomenon of suggestibility is highly reliable among children and adults, there is considerable debate concerning the mechanisms underlying the suggestibility effect. One view is that the original memory trace for the event was changed (overwritten) as a result of the suggestion. A second hypothesis is that the postevent suggestion interferes with recollection because it renders the original memory unretrievable but unchanged, as in the case of creating access competition. Whereas these first two hypotheses posit memory impairments (which reflect storage failures) as the basis of suggestibility effects, a third hypothesis is that suggestibility effects reflect gap-filling strategies rather than a memorial distortion of the original event (e.g., McCloskey & Zaragoza, 1985a); subjects accept the misleading information because they have no memory for the original event. A fourth hypothesis is that suggestibility effects result from retrieval difficulties that reflect source monitoring difficulties. According to this view, the subject has simultaneous access to representations of the original event as well as to the erroneous suggestion but has difficulty distinguishing which one *was* the original event. Source confusions might occur when only the erroneous suggestion comes to mind, that is, even when the original event cannot be retrieved (Lindsay, 1990). Source monitoring difficulties can reflect source monitoring decisions that are fast and made without conscious deliberation, or they can reflect conscious processes, such as when the subject realizes that two competing memories exist and therefore carries out a deliberate reflective analysis to determine which is the original source. Finally, some researchers (e.g., McCloskey & Zaragoza, 1985a) have posited that suggestibility effects arise out of social pressures: The subject accepts the misleading information to please the experimenter or because the experimenter is trusted. In this section, we focus on the first three hypotheses. The claims about source monitoring problems and social influences are discussed later.

In the course of encoding an event, the memorizer carries out a string of pattern recognition and interpretative analyses. The former entails the abstraction of the features of the event, such as its contrast, shape, contour, and size, whereas the latter entails attaching meaning to the event, such as naming it, assigning it an emotional valence, or forming semantic associations to it. According to trace theorists (e.g., Brainerd & Reyna, 1988; Tulving & Watkins, 1975; Zaragoza, Dahlgren, & Muench, 1992), a memory trace is the record of such pattern recognition and interpretative analyses that are carried out at the time of encoding.[3] Over time, and in response to erroneous suggestions, a trace's features may begin to loosen until they are nearly "disintegrated." At the time of retrieval, it is possible for its features to be reassembled. Thus, there are both encoding and retrieval opportunities for distortion. In addition, trace theorists (Brainerd, Reyna, Howe, & Kingma, 1990) assume that the more interpretative semantic features are less vulnerable to encoding and to retrieval manipulations of all types (e.g., delay, interference, modifiability).

Trace theorists assume that the incorporation of postevent information occurs as a function of the strength of the trace, with weak traces being especially vulnerable to featural dilution or blending (i.e., "destructive updating") or erasure (Ceci, Toglia, & Ross, 1988). Incorporation of postevent information also occurs as a function of the degree of the trace's "fuzziness," with interpretative or gistlike traces being more resistant to postevent suggestions than verbatim traces.

Two mechanisms have been invoked to account for the

[3] There are many subtle differences among trace theorists, and the description of these is beyond the scope of this article (the interested reader should consult Howe, 1991, for a description of the differences among various trace-strength approaches).

greater susceptibility of weak traces to being altered. One has to do with the nature of weak traces themselves, which tend to be loosely integrated, thus permitting greater intrusion from external sources (Brainerd, Kingma, & Howe, 1985; Brainerd et al., 1990). Because a trace is a concatenation of features that represents the original event, once these features begin to unravel it permits the incorporation of suggested features. This type of incorporation or blending of features is less likely when a trace's features are tightly bundled. On the other hand, a weak trace may provide a more hospitable encoding context for an erroneous suggestion to be admitted into memory as a *coexisting trace.* Thus, in addition to the incorporation of isolated features of the erroneous suggestion into the original trace, it is also possible that the entire erroneous trace is encoded and is allowed to attain a status on par with the original trace in terms of its strength. This happens because the contents of a weak original trace may be inaccessible at the time the erroneous suggestion is made, thereby making the intruding suggestion more likely to be subsequently recalled because there is no strong coexisting trace for the original event to compete with (Ceci et al., 1988; Howe, 1991; McCloskey & Zaragoza, 1985a; Zaragoza et al., 1992).

One important prediction of trace theory is that age differences in memory impairments will occur because younger children encode weaker traces, which are more vulnerable to featural disintegration or overwriting and also because young children encode more verbatim perceptual features and fewer interpretative or gistlike representations than older individuals. As mentioned earlier, verbatim representations should not survive as long as gist representations because they are more susceptible to disintegration. There is some support for these predictions (Toglia, 1991). Several researchers have reported that children with weaker memories of the original event are less likely to resist to suggestions about that event (King & Yuille, 1987; Warren, Hulse-Trotter, & Tubbs, 1991). Although these researchers did not directly show that the postevent suggestions overwrote an original trace, Warren et al. (1991) concluded that misinformation exerts its strongest effect on traces that were previously unreported (i.e., those traces that are presumably weakest). The converse of this position is that if a child has a strong memory trace for an event, he or she will be highly resistant to suggestion (e.g., Goodman, 1984b). Hence, it may be futile to try to overwrite a child's memory for traces that have become strengthened as a result of repetition, such as a child's name or sex.

Thus, trace theory models make a priori predictions about the conditions under which age differences in suggestibility may be pronounced or attenuated. For example, Lindberg (1991) showed that there are times when older children will actually be more suggestible than younger children, such as when the younger child's greater knowledge about some material permits stronger, gistlike encodings or when older subjects' greater knowledge leads them to make erroneous inferences that are impossible for the younger children to make because of their lack of the requisite knowledge. A good example of this latter situation can be found in the study by Duncan et al. (1982), mentioned earlier. When those researchers analyzed only the trials in which subjects correctly answered all of the control questions (i.e., the questions that were not related to the correct or incorrect postevent suggestions), they found that the youngest subjects were *less* likely to be influenced by erroneous postevent information than were the college students.[4]

However, barring the exceptional cases in which young children encode more enduring traces, the normal developmental path is from weak, verbatim traces to more durable gistlike traces. This more common pattern will ordinarily result in age-related differences favoring reduced suggestibility in older subjects.

One feature of the trace strength work that deserves special mention is the claim that some, perhaps even most, of the variance in the observed age differences in suggestibility resides at the time of *encoding* as opposed to during retrieval. On the basis of extensive analyses, some theorists conclude that age differences in suggestibility arise primarily because younger children store traces that are more apt to be overwritten by subsequent suggestions (i.e., trace destruction), not because younger children have more difficulty retrieving traces, the so-called trace competition (Lindberg, 1991). In a developmental study, Howe (1991) found that when misinformation effects occurred, they were related to alterations in the original trace and not to trace competition. If replicated and extended, this would suggest that some of the susceptibility of younger children to postevent suggestion might involve actual trace destruction as opposed to trace coexistence or competition. To the extent that this is true, subsequent probing or context reinstatement cannot undo the damage caused by erroneous suggestions.

Some trace theorists have challenged the notion that mem-

---

[4] Duncan, Whitney, and Kunen's (1982) procedure might have underestimated younger children's suggestibility in two ways: First, because it was based exclusively on trials in which the trace was strong, this might have favored immutability (Brainerd, Reyna, Howe, & Kingma, 1990). However, when Duncan et al. did not exclude weaker trace data from their analyses, age differences favoring younger children disappeared. Second, to the extent that there exist age-related differences in the contents of an encoding, the use of postevent semantic information to alter a visual memory might have worked in favor of first graders being less suggestible because they might have encoded primarily perceptual or *verbatim* features from the visually presented slides of the Duncan et al. *Star Wars* story, whereas the older subjects might have encoded the *gistlike* semantic features of the story. Hence, the use of verbal postevent information might have been encoded semantically (as gist) by the college students and integrated with their earlier semantic codes, whereas younger subjects might have encoded a verbatim trace that was never integrated with the postevent gistlike questioning. Duncan et al. acknowledged this possibility in their conclusion. This is the only study of which we are aware that has used a traditional suggestibility design and found greater suggestibility effects for adults than children. We did not describe in the text a study by Leippe, Romanczyk, and Manion (1991), which reported that adults "acquiesced" more than children, because the questions used in this study would not be considered nonsuggestive in most traditional suggestibility paradigms (see Goodman, Rudy, Bottoms, & Aman, 1990, p. 260). Moreover, the results of the Leippe et al. (1991) study can be interpreted differently depending on which questions are regarded as reflecting acquiescence. This makes it unclear if acquiescence is conceptually the same as suggestibility.

ory impairment is related to trace strength and more generally that suggestibility reflects memory impairments. The Modified Test was introduced by McCloskey and Zaragoza (1985a) as a means of determining the degree to which suggestibility effects reflect memory impairments. This test is similar to the standard paradigm developed by Loftus et al. (1978), but instead of asking subjects to choose between the original event and the erroneous event (e.g., a hammer and a wrench in the example provided earlier), they are required to choose between the original event and a new event that has not been seen or suggested (e.g., a hammer and a screwdriver in the example provided earlier). If subjects' memory for the hammer has been impaired by the provision of the wrench suggestion, then when shown the hammer and the screwdriver, subjects should select the screwdriver more often compared with a nonmisled control group. However, if subjects select the screwdriver as frequently as a nonmisled control group, then suggestibility effects reported in standard procedures (i.e., selecting the wrench over the hammer) do not reflect memory impairment but social factors (which are discussed in a separate section) or gap-filling strategies. That is, the subject does not remember the original event but does remember the postevent information and uses this to "fill in the gap" of the missing memory. The gap-filling-strategy hypothesis seems particularly relevant for the elucidation of developmental differences in children's suggestibility in that there are reliable age differences in rates of forgetting (Brainerd et al., 1985, 1990). Thus, when asked to reconstruct the original event, younger children may readily accept misinformation to fill in missing memories.

The Modified Test has produced mixed results in the study of children's suggestibility. Zaragoza et al. (1992) found no evidence of memory-based impairment in four separate experiments. Ceci et al. (1987) reported evidence of memory impairment in their third and fourth experiments, and Delamothe and Taplin (1992) reported evidence of an impairment using the Modified Test, with both 5- and 10-year-olds. There are several procedural differences between these three sets of studies (e.g., the number of times the suggestion was given, whether it was a within- or a between-subjects design, the length of the retention interval), but it is not obvious a priori why these differences should have resulted in such different outcomes.

In a slightly different paradigm in which kindergarten and Grade 2 children were told a story, provided with misleading information and then several days later asked to recall the story, Howe (1991) found that the children showed little evidence of memory impairments. That is, although children added more intrusions to their recollections when their encoding of an event was weak, the content of these intrusions was not related to the erroneous suggestions. He concluded that

> the degree of trace strength . . . is not directly related to memory impairment effects. That is, although trace strength is directly related to the rate of forgetting and the number of schema- and misinformation-relevant intrusions, it does not impair recall of the original story details. . . . Overall, . . . weak traces are no more susceptible to misleading information than strong traces. (Howe, 1991, p. 760)

This is a view echoed by Zaragoza et al. (1992).

The resolution to this dispute may have to do with the "boundary conditions" on the memory impairment effect. It may be that memory impairment occurs only when certain conditions exist, such as the strength of the erroneous suggestion (a function of, among other things, the number of times that the erroneous suggestion is made) in interaction with the strength of the original trace. Certain events, such as dynamic ones that involve actions, may have more durable trace strengths that render them more resistant to alteration than other types of events, such as those containing static attributes. For instance, Schwartz-Kenney and Goodman (1991) found that although memory-based impairment does occur for 6- and 9-year-olds, it does so only for memories having to do with person and location information, not for memories of actions. On the other hand, Rovee-Collier and Borza (in press) conducted five experiments with 3-month-olds in which evidence for memory impairment seemed strong. Infants were trained to kick in order to make an overhead mobile move across their crib. Once they acquired this association between kicking and mobile movement, they were exposed to a novel mobile overhead that was unconnected to their kicking. During later tests of memory, features of the novel mobile impaired infants' memory for the original mobile. Rovee-Collier and Borza reasoned that their paradigm contained all of the ingredients of the Modified Test: First, infants witnessed an event (a mobile moving). Second, they were exposed to postevent information that conflicted with their original memory (a new mobile that did not move in conjunction with their foot kicking). Third, they were tested for their memory of the original event.

Of course, this area of study is still in its early stages. Although the Modified Test procedure sometimes fails to produce evidence of memory impairment, these results by themselves do not indicate the source of children's (or adults') difficulties on standard misinformation paradigms; they merely indicate that acceptance of the postevent misinformation on some occasions does not reflect memory impairment. However, the results do not indicate whether social mechanisms or other cognitive mechanisms (e.g., filling-in-the-gap procedures) underpin suggestibility effects. It is also the case that the correlational data presented on the relation between children's memories and suggestibility may just as easily support a gap-filling position (no memories or unreported memories are associated with acceptance of the misleading information) as a memory impairment position. Finally, as must also be clear from this review of the literature, few of the studies were developmental. Thus, more research is required to examine the memorial bases of suggestibility effects in children as well as to determine the degree to which age-related increases in memorial skills are directly related to age-related increases in resistance to suggestive questioning.

Linguistic

Linguistic competence is also implicated in suggestibility. Because many of the studies of suggestibility include a wide age range of children of varying levels of language skill, it is reasonable to assume that there may be age differences in understanding the original events if they are verbally presented (see Garbarino & Stott, 1989; Nurcombe, 1986, for similar points). It is

also possible that children's understanding of the lexical items and syntactic structures used to test their memories may differ as a function of age. (Children's understanding of the social intents of verbal interactions are discussed in a later section on social factors.) Finally, adults may incorrectly interpret children's verbal reports as a result of children's limited production skills.

Some researchers have documented children's limited understanding of legal concepts (Saywitz, 1989; Warren-Leubecker, Tate, Hinton, & Ozbek, 1989) or have examined the complexity of the language used within the courtroom setting (Walker, 1988) and children's actual comprehension of courtroom language (Brennan & Brennan, 1988). Although these data provide important glimpses into children's comprehension of legal language, these studies do not bear directly on the issue of suggestibility.

One of the few studies to examine the effects of linguistic structure on children's suggestibility was inspired by work with adults (Loftus & Zanni, 1975). After viewing a short film, adults were asked several questions about the events. They gave more false recognitions to questions with definite articles (e.g., Did you see *the* car?) versus indefinite articles (e.g., Did you see *a* car?). The same pattern was found for 4- and 5-year-olds (Dale, Loftus, & Rathbun, 1978). The use of definite articles in questions produced more answers to questions about nondepicted events. Thus, young children appear to have the same understanding of this particular linguistic marker as do adults.

A second study examined children's responses to questions containing marked and unmarked modifiers. An example of a pair of unmarked and marked modifiers is *fast* versus *slow*. The unmarked term *fast* carries no assumption concerning an upper or lower bound, whereas the marked term *slow* implies the absence of the property in question. Children generally acquired the unmarked form of a pair before the marked form (Clark & Clark, 1977). Lipscomb, Bregman, and McCallister (1984) showed first- through eighth-grade children and college students a film of an automobile collision. Subjects were then questioned about the speed of the car using marked or unmarked terms (e.g., "How slow/fast was the car going when it smashed/hit the other car?"). Subjects of all ages provided faster estimates to questions with the word *fast* than to questions with the word *slow*. However, only seventh and eighth graders provided faster speed estimates when the word *smashed* was used. These results are suggestive only of developmental differences because the adult subjects did not provide different estimates for *smashed* and *hit*, a result discrepant with that obtained by Loftus and Palmer (1974).

Hence, it is possible that the way questions are worded may affect age patterns in suggestibility. If young children do not have full syntactic or semantic understanding of linguistic features, they may not be biased in the same way as older children, thus canceling out any potential suggestibility effect. Therefore, the inconsistent results of some developmental studies may reflect subtle linguistic differences among stimuli.

Knowledge

Semantic. Semantic knowledge refers to an individual's repository of world knowledge about the declarative, procedural, and associative meanings of concepts. Memories of events reflect, among other things, how much was known about the event prior to its observation and how this knowledge is represented in memory. Thus, chess masters recall board positions of past games better than nonmasters, and baseball experts remember more details about baseball games than nonexperts (see Chi & Ceci, 1987). Occasionally, younger children possess superior knowledge to older children (e.g., about cartoons), and in these cases they often excel at remembering (Chi & Ceci, 1987; Lindberg, 1980).

In addition to the sheer amount of factual knowledge, the representation of this knowledge in long-term memory (i.e., its relational and implicational structure) plays an important role in recall. If a 7-year-old's knowledge about a character is that she is strong and smart, then an implication might be that she is also attractive. When children with such beliefs try to recall incongruous events about characters (e.g., a character is smart and strong yet unattractive), they often can do so if tested immediately, while the trace is still strong. However, when the trace has been weakened because of a 3-week delay, children make erroneous recalls that are consistent with their prior knowledge but inconsistent with what they actually saw (Ceci, Caves, & Howe, 1981). Similarly, 5- and 6-year-olds who have differentiated sex-based knowledge often incorrectly select the pictures seen 1 week previously of male and female actors performing sex-incongruent tasks (Martin & Halverson, 1983). They report having seen a picture of a male actor playing a traditional male role even though it was a female actor who was depicted in the picture.

Thus, developmental differences in the structure of semantic knowledge can lead to different inferences about witnessed events. Usually, increased knowledge facilitates recall, but not invariably. One qualification to this broad conclusion is warranted: When the event being recalled is so unlike a child's representation that it appears bizarre, then it is recalled more accurately (Davidson, 1991). Nonplausible details and events can actually facilitate memory if they are so different from the child's knowledge as to appear bizarre.

Scripted knowledge. Temporally organized, habitual, agent–actor–action routines are referred to as *scripts*. For example, a restaurant script includes the expectation that the customer is first seated by the waitress, given a menu, places his or her order, and so forth. Scripts serve to generate expectations, and when the expectations run counter to what occurred, the result can be that scripts produce an erroneous reconstruction of the events.

Although scripts develop with age, even very young children possess these for familiar events. These scripts influence the child's reconstruction of previously experienced events. In certain conditions scripted knowledge may exert positive effects on memory reconstruction. For example, if children have a script for the sequence of events that unfolds in a normal school day, they may unconsciously use scripts to fill in gaps when their actual memory has faded (Myles-Worsley, Cromer, & Dodd, 1986).

However, scripted knowledge may also exert negative effects on memory. If a girl has attended multiple gym classes, she is more likely than a child who has attended only one workshop to

erroneously claim that a particular habitual act (e.g., stretching) occurred even when it did not (Hudson, 1990). This is because attending multiple gym classes or school events that share the same structure leads to the creation of generalized scripts, something that one-time attendees do not possess.

The relation between scripted knowledge and accurate recall may change as a function of age, depending on the level of children's scripted knowledge and the characteristics of the to-be-remembered event (Ceci et al., 1981). Once children have acquired scripts, preschoolers' recall may be more vulnerable to the negative effects of script-based knowledge than elementary school-aged children. Hudson and Nelson (1986) summarized their research in this area by concluding that preschoolers were less flexible than older children; they were more likely to read off scripts than to recall single episodes. When the information was discrepant or unexpected in relation to scripts, preschoolers had more trouble recalling this information than older children. Farrar and Goodman (1990) elaborated this position by examining recall in relation to children's development of scripted knowledge. Young (4-year-old) and older (7-year-old) children experienced an unfamiliar event in a laboratory on several different occasions. These repeated experiences were intended to allow the children to develop script-based knowledge for these "standard" events. The children then experienced a novel set of events called *deviation events.* One week later, the children were asked to recall the standard and deviation events. The older children were able to distinguish between their memories for the standard and deviation events. That is, they developed a script for the standard events, and the deviation events were separately tagged in memory. By contrast, the younger children confused the events from the two different sets of experiences; they incorporated the deviation events into their developing schemata of the standard events. The results of these studies provide a theoretical basis for age differences in suggestibility. Younger children are more suggestible because they are overly dependent on scripted knowledge and incorporate discrepant or novel events (which could be a suggestion) into their script of the event rather than keeping them tagged as separate events.

When younger children's scripted knowledge is insufficient or poorer than that of older children, older children might make more false inferences about events that are not witnessed but that are part of their scripts. Lindberg (1991) showed this to be the case. Because of their more elaborate scripts about how cheating could occur, sixth graders and college students made more false attributions than third graders about an ambiguous event. When subjects were erroneously told that the film they were viewing depicted cheaters, sixth graders and college students tended to report cheating that was based on innocent acts such as a student asking another for the time of day. Younger children's scripts for cheating did not contain this scenario as a pretext for cheating, so their limited script knowledge made them less prone to the erroneous suggestion. Along the same lines, adults are more likely than young children to assume that on meeting someone, they are to shake the person's hand because doing so is part of their script for new encounters (Goodman & Reed, 1986).

The finding of Duncan et al. (1982) of greater suggestibility on the part of older subjects may also reflect the interfering effects of scripted knowledge. Their task required the integration of high levels of scripted knowledge, which the youngest children probably did not possess, thereby precluding its effectiveness as a source of biasing. For example, the supposition embedded in the question, "Was the hunter's fishing pole broken by a bear?" (in reality, the hunter did not have a fishing pole but a spear), may be more easily integrated into a college student's "hunting" script than a first grader's, thus leading adults to integrate the misinformation with the original information more readily than younger children.

In summary, it seems reasonable to assume that a positive relation exists between the amount and structure of knowledge and children's resistance to suggestion, at least in cases in which an event is congruent with existing knowledge. In other situations, older children and adults may be more suggestible than young children because their greater knowledge might lead them to infer script-relevant details that were omitted from the actual event or to integrate postevent information with the original event.

Stereotypical knowledge. Stereotypes are naive theories about personal characteristics that organize and structure experience by directing individuals to look for certain types of information and advising them on how to interpret it. Stereotypes are a form of schematic knowledge that helps organize memory, sometimes distorting what is perceived by adding thematically congruent information that was not perceived (Martin & Halverson, 1983).

Little is known about whether there are reliable age differences in the tendency to extrapolate from stereotypical knowledge to provide erroneous but plausible accounts of nonwitnessed events. However, even 3- and 4-year-olds will sometimes be misled and claim to have witnessed events that did not occur but that are consistent with a stereotype. For instance, a character named Sam Stone was described over a 1-month period to 3- to 6-year-olds as someone who was very clumsy and who always broke things that did not belong to him (Ceci, Leichtman, & White, in press). After this stereotype-induction procedure, Sam Stone visited the children's nursery school, where he spent 2 min amiably interacting with the children during a group story-telling session. At that time he did not behave clumsily or break anything. The following day, the children were shown a ripped book and a soiled teddy bear. They were asked if they knew how the book had been ripped and the teddy bear soiled. Few children claimed to have seen Sam Stone do these things, but 25% said that *perhaps* he had done it—a reasonable statement given the stereotype-induction they had received. Next, the children were interviewed once per 2 weeks for 2 min each over the course of the next 10 weeks. During each interview, the children were asked two leading questions such as "I wonder whether Sam Stone was wearing long pants or short pants when he ripped the book?" or "I wonder if Sam Stone got the teddy bear dirty on purpose or by accident?" These suggestive questions were consistent with the stereotype that had been previously provided, and nearly all of the children answered them. At the end of this 10-week interrogation period, the children were interviewed by a new interviewer who told them she was not at their school the day Sam Stone visited and wanted to

know everything that happened. When asked, 72% of the 3- and 4-year-olds said Sam Stone had ruined at least one of the items in question (the book or bear). When they were explicitly asked, 45% of the 3- and 4-year-olds replied that they actually had seen him do these things, as opposed to merely being told that he did. These false accounts often were embellished with perceptual details (e.g., Sam Stone took a paint brush and painted melted chocolate on the teddy bear, Sam took the book into the bathroom and soaked it in warm water until it fell apart) or emotional details (e.g., Sam was acting very silly when he spilled coffee on the bear, Sam was mad and ripped the book with his hands). In contrast to the 3- and 4-year-olds, only 11% of the 5- to 6-year-olds claimed to have actually observed Sam Stone damage the items. A control group, who received only the multiple suggestive interviews, with no prior stereotypical knowledge about Sam, made significantly fewer false claims than children who were given stereotypical knowledge. Thus, these results indicate that not only do young children form stereotypes but that stereotype formation interacts with suggestive questioning to a greater extent for younger than older children.

Much work remains to be carried out on the relation of semantic, scripted, and stereotypical knowledge to suggestibility. Nevertheless, the existing work indicates that the quality and quantity of memory representations influence subsequent recall and susceptibility to suggestibility. Although most of the time this works in favor of older children and adults, special circumstances can be found wherein younger children's lack of knowledge actually prevents them from succumbing to an erroneous suggestion.

Source Monitoring: Distinguishing Reality From Fantasy

An important but relatively unexplored cognitive variable is the extent to which suggestibility in children arises from an incapacity to distinguish between the various sources of their memory. Freud (1933/1966) postulated that claims of childhood sexual abuse by his female adult patients were false, reflecting their inability as children to distinguish between reality and fantasy (however, see Masson, 1984, for an alternate account). Freud thought it possible to retrieve original memories through the removal of symbolic transformations that "blockaded" them from consciousness. Piaget (1926), however, was less optimistic that early memories could be separated from fantasies, commenting that "the child's mind is full of these 'ludistic' (fantasy play) tendencies up to the age of seven or eight, which means before that age it is very difficult for him to distinguish the truth" (p. 34).

Outside of the classical work on animism by Piagetians, the topic of reality monitoring did not receive empirical scrutiny until the 1970s, when a number of researchers converged on the view that young children were able to distinguish between reality and fantasy (Flavell, Flavell, & Green, 1987; Morison & Gardner, 1978; Taylor & Howell, 1973). For example, Morison and Gardner (1978) presented 5- to 12-year-olds with three toys and asked them to group fantasy figures (e.g., a dragon and elf) and to exclude real figures (e.g., a frog). Even 5-year-olds were highly accurate, although errors decreased with age.

Harris, Brown, Marriott, Whittall, and Harmer's (1991) results modify these conclusions in important ways. As in previous studies, 4- and 6-year-olds reliably distinguished between fantasy and reality; most children rightfully stated that imagined ghosts, monsters, and witches were not real. However, when told to imagine a pretend character that was sitting in a box, after a short period of time many of the children began to act as though the pretend character was real. For example, half of the children were told that the pretend character was a rabbit and half were told that it was a monster. The experimenter then said she had to leave the room for a few minutes; four of the twelve 4-year-olds who were told that there was a pretend monster in the box would not let her leave the room, even though they had just seen and stated that the box was empty. None of the other children acted this way. When the experimenter returned, almost half of the children in both age groups wondered whether perhaps there was an imaginary creature in the box. Questioning uncovered some magical and unrealistic thinking. Although almost all of the children admitted to pretense before the experimenter's departure, 25% of the children now thought that pretend creatures could become real. These data reflect the fragile boundaries of children's fantasy–reality distinctions. When situations become intense, children appear to easily give up these distinctions. Although the children were repeatedly assured that the creatures were imagined, it seems that the experimental procedures were highly suggestive, breaking down their shaky differentiations. The 4-year-olds had more fragile boundaries than did 6-year-olds. In both case studies (summarized at the beginning of the article), children's disclosures became increasingly fantastic and incredible. This escalation could have been caused by interviewers not drawing the children back to reality when they made "fantastic" claims and perhaps as a result, their perceptions moved from fantasy to reality for the children, leading them to believe themselves.

Another area of research focusing on later detail on young children's difficulty distinguishing between what they experienced through perception and what they only imagined they experienced. Johnson and her colleagues have been at the forefront of this area for a decade (see M. K. Johnson, 1991, for a review; Foley & Johnson, 1985; Foley et al., 1991). In the most comprehensive model, the "multiple-entry modular memory system," recollection is based on the interplay of two subsystems. The perceptual system records and stores the contents of perceptual processes such as seeing and hearing, whereas the reflective system records psychologically generated information such as imagining, thinking, and speculating. Developmental differences in reality–fantasy monitoring could reflect the earlier functional capability of the perceptual subsystems and the later development of the reflective systems. At issue is whether these subsystems are developmentally invariant or unfold over a long period of time (Lindsay et al., 1991).

When asked to judge whether they had said a word versus imagined saying it, 6-year-olds have more difficulty discriminating between these two sources of memories than 9-year-olds and adults (e.g., Foley et al., 1983). Apparently, the cues involved in differentiating between certain types of actual and imagined events may not be well developed before late childhood. Be-

cause young children do not have difficulty distinguishing between something they said (or did) and something that someone else said (or did), it seems that they can differentiate between these sources of their memories except in situations in which at least one of the sources is self-generated (Foley, Santini, & Sopasakis, in press). Specifically, younger children are more error prone at distinguishing between real versus imagined acts or words when both concern themselves, but they are no worse than adults when it comes to judging whether an act (or words) was performed (or spoken) or imagined by themselves versus someone else.

Recently, however, a more general source monitoring framework has been invoked to account for young children's source confusions. According to this account, young children find it especially difficult to separate sources of information that are perceptually and semantically similar. For example, 7- and 10-year-olds and adults were shown a videotape of a set of actions and were instructed to either perform, to watch others perform, to imagine themselves perform, or to imagine another perform these actions (e.g., "Please watch the girl touch her nose" vs. "Please imagine touching your nose"). Subjects were later asked to indicate for each of a list of actions which acts had actually been performed and which they had imagined and which were new. Compared with adults, children found it more difficult to distinguish between imagined and actual actions if the same actor was involved in both kinds of actions (e.g., watching vs. imagining the girl touch her nose). By contrast, young children performed as well as adults when the sources of information were relatively discriminable (self vs. girl). Thus, although all age groups reliably distinguished between the actions of two perceptually or semantically distinct actors, "children are more likely to confuse memories from different sources whenever those sources are highly similar to one another" (Lindsay et al., 1991, p. 18).

Source monitoring studies suggest that children could be vulnerable to a range of confusions between actual events and suggested events when they are perceptually and semantically similar. However, because the locus of children's greater misattributions is unclear, and there are no data that link children's suggestibility and source monitoring difficulties (see Lindsay, 1990, for adult data), these claims are speculative at this stage.

Summary: Cognitive Abilities

Although our review of cognitive factors does not include all of the cognitive variables that could conceivably be involved in age differences in suggestibility (e.g., inferential skills; abstract reasoning abilities; perspective-taking, metacognitive skills), it does describe the factors that have received the attention of researchers. In view of the previous discussion of the fundamental role that the development of these abilities plays in decreasing children's susceptibility to suggestion, it is not surprising that IQ tests that measure many of these cognitive skills correlate with children's levels of suggestibility (e.g., Hurlock, 1930; Otis, 1924). Recently, Haugaard and Repucci (1992) reported that although IQ was unrelated to children's accuracy in realizing that another child was inaccurate in claiming that her neighbor hit her, preschoolers with low IQ scores were more likely to erroneously attribute this inaccuracy to truthfulness on the part of the child. We now discuss social and motivational factors that need to be considered to explain age differences in suggestibility.

Causal Mechanisms: Social and Motivational Factors

As originally suggested by early researchers such as Binet, Stern, Varendonck, and Lipmann, children's suggestibility is not purely a cognitive phenomenon but also reflects social and motivational factors. After a 70-year hiatus, modern researchers have begun to examine the potential effects of selected social and motivational factors on children's suggestibility. In this section, we focus on specific conditions within an interview that induce compliance to the interviewer's misleading questions.

An interview is successful when the interviewer obtains a complete and accurate account from the interviewee. In order to achieve this goal, more is required than the accurate comprehension and production of linguistic utterances. In addition, participants must understand a broader set of conversational rules that bind the questions and answers. Of particular importance to the present topic is the degree to which children's performance in an interview reflects their understanding of the social rules underlying conversations.

In an interview, the listener tries to figure out the speaker's intent; often, this involves going beyond the direct meaning and computing an indirect meaning. The number of interpretations of messages, however, is constrained by the social conventions and context of the interview. These social conventions include the "principle of cooperativity" (Grice, 1975), which states that listeners interpret speakers' utterances on the assumption that they are informative, true, relevant, and clear. These assumptions about cooperativity, which are used to infer the meaning and intent of utterances, may change as a function of the social relationships, perceived motivations, beliefs of the participants, and the actual setting of the conversation. Good listeners ask, "Given the context of this conversation, what is the intended meaning of the utterance?" If there is a disjunction between a questioner's goals and a listener's perceptions of these goals, the interaction will not be successful.

From an early age, children perceive their adult conversational partners as being cooperative, truthful, and not deceptive (Garvey, 1984; Nelson & Gruendel, 1979; Romaine, 1984). Children are also cooperative partners; they supply their adult questioner with the type of information they think is being requested (e.g., Ervin-Tripp, 1978; Read & Cherry, 1978). This pattern reflects children's desire to comply with a respected authority figure. As a result, when questioned by adults, children sometimes attempt to make their answers consistent with what they see as the intent of the questioner rather than consistent with their knowledge of the event. Several pieces of data support these contentions.

First, young children perceive adults as being highly credible and competent sources of information; they place more faith in the credibility of adults' statements than in those of their peers. Sonnenschein and Whitehurst (1980) reported that 6-year-olds became better referential speakers after listening to competent

peers, but not after listening to incompetent adults, because they assumed that all adults are competent. Ackerman (1983) presented first and third graders and adults with paragraphs containing contradictory information between a contextual source that was either authoritative (e.g., a doctor) or nonauthoritative (e.g., a janitor) and a speaker that was either a child or an adult. Subjects were asked which source of information was the most believable. Adults based their judgments on the authority of the source, whereas first graders based their judgments on the age of the speaker, with adult speakers being rated as more believable than children regardless of authoritativeness. Third graders weighed both sources of information. These data suggest that young children are biased to believe adults and to accept their statements as credible.

Second, children attempt to answer adults' questions even if the questions are bizarre. When asked nonsensical questions such as "Is milk bigger than water?", most 5- and 7-year-olds replied "yes" or "no"; they rarely responded "I don't know" (Hughes & Grieve, 1980). These data suggest that children perceive adults as being cooperative conversationalists who ask honest and logical questions that must have an answer. Pratt (1991) reported that adults are not immune to such pressures; they (i.e., adults) also provide answers to some types of bizarre questions even if preschoolers are more pervasively willing to do so.

Finally, when children are asked the same question more than once, they often change their answers presumably because they interpret the repeated question as "I must not have given the correct response the first time; therefore, to comply and be a good conversational partner, I must try to provide new information." In Cassel and Bjorklund's (1992) study of children's memory for a bicycle theft, 42% of kindergarten children changed their mind on repeated questioning. Young children's responses to Piagetian conservation questions are more accurate when they are asked only once versus several times, as a result of their proclivity to change their answers in response to their impression of what the interviewer wants them to say (Gelman, Meck, & Merkin, 1986; Rose & Blank, 1974; Siegal, Waters, & Dinwiddy, 1988). Siegal et al. (1988) showed young children a videotape of a puppet being given a conservation test. After the puppet made a response, the children were asked if it had responded to please the adult or because that was what the puppet really thought was true. The children were more likely to say that the puppet pleased the experimenter when he gave an incorrect response in a two-question interview. By contrast, they were more likely to say that the puppet really thought the answer was true in the one-question procedure.

Although these results suggest that repeated questioning within the same session may decrease the consistency of children's performance, they do not address the potential effects of repeated questioning across (as opposed to within) sessions. Repeated questioning across sessions has at times facilitated memory, possibly because recalling an event is a form of rehearsal that serves to reactivate traces. This has been found in several studies in which children have recalled approximately 10% more information on repeated recall versus a single test (Baker-Ward, Hess, & Flannagan, 1990; Brainerd et al., 1990; Tucker, Merton, & Luszcz, 1990). In other studies, however, facilitation has not been found (Dent & Stephenson, 1979; Flin et al., 1992; Ornstein et al., 1992; Steward, 1989; Warren & Swartwood, in press). For example, Dent and Stephenson (1979) found that 10- and 11-year-olds gave virtually identical answers to the same questions that were posed on 2 consecutive days. Similarly, Steward (1989) found that repeated questioning of children between the ages of 3 and 7 years led to equivalent performance regardless of whether the children were questioned one, two, or three times over a 6-month period.

The data just reviewed about social factors are relevant to the issues of children's testimony and suggestibility because they point to the possibility that children provide incorrect answers to suggestive questions because they view the interviewer as asking credible questions and thus incorporate the content of the question into their answers. Unlike adults, children may rarely challenge the credibility of adult questioners. If true, then the following manipulations to an interview should result in reduced suggestibility. When children rather than adults provide misleading information to other children, suggestibility effects should be diminished because peers should be viewed as being less trustworthy and authoritative than adults. Also, if children are told that it is permissible to say "I don't know" or if they are warned that the experimenter may be trying to trick them, they should be less suggestible. Finally, if children are asked repeatedly to recall a specific event, their responses to the first question should be more accurate than their responses to repeated questions because the first question does not imply that a prior answer was incorrect or otherwise undesirable. Data supporting some of these hypotheses exist.

Ceci et al. (1987) presented short stories accompanied by illustrations to preschool children. One day following the presentation, an adult (Experiment 1) or a 7-year-old child (Experiment 2) provided misleading information about certain aspects of the stories. Two days later, the children were questioned. Children were less suggestible when they were given the same misleading information by a child than by an adult, indicating that they yielded the contents of their own memory to that of a more powerful adult authority figure—something they did less often when the misinformation was supplied by another child.

Moston (1987) questioned 6- to 10-year-olds about an event that was observed during a school assembly. Half the children were told that it was acceptable to give "I don't know" responses. Provision of the "don't know" instruction did not result in increased accuracy or in "don't know" responses. Contrary to other results, even children in the control group gave a number of "don't know" responses, and all children gave more "don't know" responses to misleading than to direct questions. Thus, the results of this study do not provide support for the hypothesis that children are unwilling to provide "don't know" answers, especially in response to misleading questions.

When children are given some warning about the potential for deceit or false suggestions by their questioner, they are more resistant to misleading questions. Warren et al. (1991) read a story to subjects (aged 7, 12, and adult) and then asked questions about the story, some of which were misleading. At the onset of the experiment, half of the subjects were warned that the questions were difficult or tricky and that they should try to answer only what they really remembered. At all ages, warned subjects

correctly answered more misleading questions than unwarned subjects. It should be noted, however, that the effect was equally small across all age groups; the warning increased subjects' accuracy by approximately 5%.

The effect of repeated questions on children's recall has been the focus of several studies. In the Warren et al. (1991) study, subjects were told after answering the first round of questions that they did not do so well and that they should try again. When told this, children changed their answers more often than did adults. As part of Moston's (1987) study, subjects were asked the same questions twice within the same interview session. The number of correct responses significantly declined from the first question to the second question, in line with the Piagetian studies mentioned earlier. Although Moston found that accurate responses dropped overall, from 69% to 54%, the effect of repeating a question was especially dramatic for the youngest children (6-year-olds), whose accuracy fell from 60% to 39%. In contrast to their 21% drop, the two older groups' accuracy rate dropped 9%–16%. Moston interpreted the decline to have been caused by the children's belief that the experimenter was "telling" them that their first answer was wrong or unacceptable.

Poole and White (1991) examined the effects of repeated questioning within and across sessions. In this study, 4-, 6-, and 8-year-olds as well as adults witnessed an ambiguous event. Half of the subjects were interviewed immediately after the event and 1 week later. The remaining subjects were interviewed only once, 1 week after the event. Within each session, all questions were asked three times. Although Poole and White did not use leading questions, their repeated use of yes–no questions can be viewed as a subtle form of suggestion. As noted earlier, simply repeating a yes–no question could have the effect of suggesting to a child that the interviewer is unsatisfied with the initial answer.

Poole and White (1991) found that repeated questioning with open-ended questions, both within and across sessions, had little effect (positive or negative) on children's or adults' responses. However, on repeated yes–no questions, 4-year-olds were most likely to change their responses, both within and across sessions. Thus, the major finding of this study was that repeated questioning may affect very young children's responses to specific questions. Although repeating open-ended questions may merely signal a request for additional information, repeating specific questions that have a limited pool of responses (yes or no) may signal to young children that their first response was unacceptable. This finding is important because young children tend to give limited responses to open-ended questions, and therefore interviewers often resort to specific questions to elicit additional information. In order to confirm a child's answer, interviewers frequently repeat the question. In a 2-year follow-up study, Poole (in press) found that the youngest children were significantly less accurate between and within interviews.

The results of these modern studies are reminiscent of those of Binet, Lipmann, and Stern, all of whom spoke of the authority of the interviewer in the eyes of the child. According to their accounts, children view interviewers' questions as imperatives to answer, or else they attempt to revise or fill in memory gaps in order to please the experimenter. This is illustrated dramatically in some court cases. For example, in a highly publicized sexual abuse case in Jordan, Minnesota, in 1984, one child later confessed that he fabricated detailed stories of abuse because "I could tell what they wanted me to say by the way they asked the questions" (Benedek & Schetky, 1987, p. 915).

Although the studies just described highlight how particular aspects of an interview may influence children's reports, these experimental settings are pale versions of interviews carried out in legal settings (McGough, in press). In the latter context, children are questioned, on average, 11 times prior to testifying in court, often by a number of different interviewers (e.g., parents, police, therapists, child protection workers, lawyers) who usually do not have a specific set of written questions. Rather, interviewers generally use a variety of on-line strategies before and during the interview to obtain the most detailed and accurate information about events that a child might have witnessed.

Interview strategies are characterized not only by the types of questions asked (open ended vs. yes–no) but also by the emotional tone or disposition of the interviewer. It is also the case that interviewers often do not have complete or accurate information about the target events that are the basis of the interview. This is important because an absence of knowledge or incorrect knowledge about the target events may affect the style of the interview, which may affect children's suggestibility. Furthermore, some interviewers may have strong vested interests in a particular type of report.

Clinical psychologists place particular importance on building rapport with young clients so that they will relaxed and nonthreatened. To achieve this goal, they act positively toward the children by encouraging and reinforcing their answers and, on rare occasion, chastising their failure to disclose. Interviewers are sometimes criticized for reinforcing and encouraging children's responses (Raskin & Yuille, 1989). It is claimed that these strategies are not conducive to accurate reporting, as can be illustrated by a social worker's interview of Child 5C in the Kelly Michaels' Wee Care case study:

> Do you want to sit on my lap? Come here. I am so proud of you. I love big girls like you that tell me what happened—that aren't afraid because I am here to protect you. . . . You got such pretty eyes. . . . I'm jealous, I'm too old for you. (7/3/85 at 12 in Point VII of the Appellate Court Brief)

A few studies have examined how the emotional tone of an interview influences children's recall, and these results are not entirely consistent. Goodman et al. (1990) interviewed 3- to 7-year-olds 2–4 weeks after they had received an inoculation. They were interviewed either by a "nice" person who gave them considerable support for their performance throughout the interview or by a neutral experimenter who provided little support for their performance by maintaining a detached manner of interacting. All children were asked a series of specific and misleading questions; in addition, they were asked a set of misleading abuse questions (e.g., "Did he take you into the bathroom?"). Data were reported only for the last set of questions. Although interview style had no effect on older children's false reports, younger children gave fewer false reports to "nice"

interviewers, and in one subclass of leading questions, the age effect was eliminated entirely. Goodman et al. concluded that young children were most resistant to misleading abuse questions, and less likely to falsely claim their clothes had been removed, or their bottoms touched, when they felt comfortable with the interviewer.

In another study (Saywitz, Geiselman, & Bornstein, 1992), detectives from a sheriff's office used a variety of interviewing techniques and personal styles when questioning third and sixth graders about a staged event. Children who were engaged in the most rapport-building events before the interview produced the fewest errors. Collapsing across various interview conditions, children questioned by unenthusiastic, neutral detectives produced the fewest accurate details but also the fewest inaccurate statements. Children interviewed by condescending detectives, who purported to have little faith in children's answers, produced more accurate statements, but also more inaccurate statements, than children interviewed by unenthusiastic detectives. Finally, those interviewed by positive detectives, who were supportive throughout the interview, produced the most accurate details; however, they also produced as many incorrect details as children interviewed by condescending detectives.

Thus, whereas positive interviewers may elicit the most accurate details from children, results of one study suggest that they also tend to elicit more inaccurate statements than neutral interviewers. These data appear inconsistent with those reported by Goodman et al. (1990). However, it is important to note that Goodman et al. reported data only for "abuse" questions. It is not known how their children responded to nice and neutral interviewers' nonabuse questions.

In another study, 3- and 6-year-old children played with an unfamiliar male for 5 min while seated across the table from him. Four years later, Goodman, Wilson, Hazan, and Reed (1989; also described in Goodman & Clarke-Stewart, 1991) reinterviewed the children. At this time, the researchers created "an atmosphere of accusation" by telling the children that they were to be questioned about an important event and by saying things such as "Are you afraid to tell?" "You'll feel better once you've told." Although few of the children remembered the original event from 4 years earlier, their performance on suggestive abuse questions was, according to the researchers, "mixed." Five of the 15 children incorrectly agreed with the interviewer's suggestive question that they had been hugged or kissed by the confederate, 2 of the 15 agreed that they had had their picture taken in the bathroom, and 1 child agreed that she had been given a bath. These data suggest that children can be influenced by an interviewer's tone and urgency to make erroneous claims about events for which they have *no* memory. Although Goodman found reason for optimism in these results (noting that none of the children claimed that their clothes had been removed or that they had been touched in a bad way or spanked), the results are damaging to the claim that children cannot be led by suggestive questions to make abuse-related claims. Furthermore, if those children were subjected to the kinds of prolonged and pressurized interviews that we describe later, it is conceivable that even more might eventually have alleged that they had been bathed or kissed in the bathtub. If a boy had told his parent that a babysitter took his picture or kissed him while in the bathroom, this could prompt the parent to pursue an aggressive and persistent line of questioning.

In a follow-up to the aforementioned study, Goodman (1990) found that 8% of college students and child protection workers who were shown videotaped interviews of the children's answers to the suggestive questions said that sexual abuse was "very likely," and an additional 10% said that abuse was "likely." Goodman (1990) was encouraged by the fact that only 18% of the adults felt that abuse had occurred, remarking that "lay persons, at least the ones in our study, were unlikely to think that the non-abused children had been abused. Had the same information been presented at a trial, it is likely that the (innocent) defendant) would have gone free" (Freiberg, 1990, p. 32). This optimistic interpretation ignores three potential outcomes that could result in unjust prosecution. First, a single juror, highly convinced that abuse occurred, may be able to persuade the other 11 jurors who are less certain about its occurrence, just as it only takes a single juror who is convinced that no abuse occurred to persuade an entire jury that it has not. If a single juror can do this, then two jurors (16.67%) will have an even easier time. Second, if a boy alleges that a babysitter or stranger took his picture or kissed him in the bathroom, this could launch an investigation that, even if it did not result in a conviction, might still be personally devastating to the accused. Third, when these data are extrapolated to forensic settings, there is even less basis for optimism because the typical forensic case would have involved *multiple* prior attempts to create an "atmosphere of accusation," not just a single one several years after the event. If children will claim to have been kissed and photographed by strangers after a single enjoinder that "they will feel better once they tell," then repeated and stronger enjoinders (e.g., "Don't you want to help us keep him in jail?") may result in greater numbers of children making similar claims. The use of repeated atmospheres of accusation is exemplified in the Country Walk Baby Sitting Service case by the following interaction between a psychologist and one of the children.

> *Dr. Braga:* Did they [Frank and Iliana Furster] ever tell you at any time that anything would happen to you if you told the secret? . . . You see, if you tell us, then it will go away and you won't have to be scared any more.
>
> *Child:* I know, um.
>
> *Dr. Braga:* We can make it go away if you just tell us anything that they told you would scare us. By telling us, it will never be anything to worry about any more.
>
> (Hollingsworth, 1986, p. 69)

Usually, interviewers have varying amounts of information about the events under question. Pettit, Fegan, and Howie (1990) examined how an interviewer's information about events would affect the style of questioning and the accuracy of the child's reports. Three- to five-year-olds participated in a staged event and were questioned 2 weeks later. Some interviewers were given full, accurate knowledge of the event; some were given a report containing inaccurate information; and others were given no information about the event. All interviewers were told to question each child until they found out what happened and to avoid the use of leading questions.

Collapsing across interview groups, the children were asked an average of 50 questions during the 20- to 30-min interview;

thus, they were put under a great deal of pressure to provide information. Despite the warning to avoid leading questions, 30% of all questions were leading, and half of these were misleading. Interviewers with inaccurate knowledge asked four to five times as many misleading questions as the other interviewers. Overall, children agreed with 41% of the misleading questions, and children who were interviewed by misled interviewers gave the most inaccurate information. Interviewers with no knowledge showed marked rises in their use of leading questions as additional children were interviewed; these interviewers extracted more inaccurate information from the children on later compared with earlier interviews. These results suggest that interviewers' knowledge influences their style of questioning, which in turn affects the accuracy of children's testimony.

It may be that inaccurate information is detrimental only when the interviewer is a stranger. When parents were given inaccurate information about an event, they were still able to elicit accurate information from their preschoolers (Goodman, Sharma, Golden, & Thomas, 1991). Replication of this result would provide an assurance of the reliability of children's disclosures to parents, as opposed to unfamiliar law-enforcement officials.

As we have seen, child witnesses are often questioned about events that may have several interpretations, at least for the child. In the legal setting, children are interviewed many times by a variety of interviewers before they ever testify in court. What happens when children are repeatedly questioned about an event that has different interpretations for different interviewers? On the basis of the data reviewed so far, one might expect young children to be most inconsistent and suggestible in this situation. The results of a study conducted by Clarke-Stewart, Thompson, and Lepore (1989; also reported in Goodman and Clarke-Stewart, 1991) support this hypothesis. In that study, 5- and 6-year-olds interacted with a confederate posing as a janitor who followed one of two scripts. In both scripts, the confederate, named Chester, cleaned the room and then began either cleaning the toys, including a doll, or handling the doll roughly and suggestively. Chester's dialogue reinforced the idea that he was either cleaning the doll or playing with it in a rough manner. The child was then questioned about this event several times by different interviewers who differed in their interpretations of the event. Their style of questioning mirrored their interpretations. The interviewer was either (a) accusatory in tone (suggesting that Chester had been inappropriately playing with the toys instead of working); (b) exculpatory in tone (suggesting that Chester was just cleaning the toys and not playing); or (c) neutral and nonsuggestive in tone. In the first two types of interviews, the questions changed from mildly to strongly suggestive as the interview progressed. Following this interview, each child was interrogated by a second interviewer who either reinforced or contradicted the first interviewer.

When questioned by the neutral interviewer or by an interviewer whose interrogations were consistent with what the child had witnessed, children's accounts were factually correct. However, when the first interviewer contradicted the script, children's stories quickly conformed to the suggestions of the interviewer; by the end of the interview, 75% of the children's remarks were consistent with the examiners' script, and 90% answered the interpretative questions in agreement with the interviewer, as opposed to what had actually happened. When questioned by parents immediately following the interview and 1-week later, children's answers reflected the interviewers' interpretation of the events.

When the second interviewer contradicted the first interviewer, the majority of children fit their stories to the suggestions of the second interviewer. Moreover, children's subsequent reports to their parents reflected a mixture of both interviewer's interpretations.

These results concerning children's reports to their parents are inconsistent with the Goodman, Sharma, et al. (1991) finding that children make accurate disclosures to their parents. Perhaps accurate disclosure to parents is obtained when the child is questioned only by the parent on one occasion.

The aforementioned studies show how emotional tone and interviewer beliefs mold the linguistic interactions of an interview and how these molded interactions may at times promote children's false reports. Because many of the authors of these studies did not make developmental comparisons, one cannot firmly conclude that these factors uniquely affect children's reports. However, it does seem plausible that adults would not be as affected by interviewer style. This hypothesis is based on other developmental evidence, presented earlier, that suggests that compared with adults, children do view adult questioners as being more authoritative and trustworthy and thus are more likely to comply with their intended requests (Moston, 1987; Poole & White, 1991; Pratt, 1991; Warren et al., 1991).

As studies on the effects of interviewing techniques become more realistic (e.g., by providing additional interviews, with interviewers differing in style), it seems likely that they will lead to the conclusion that earlier studies of the suggestibility of children's memories for neutral events in a single interview might have underestimated young children's suggestibility in real-life interviews. However, they also might have overestimated the amount of suggestibility that arises from interviews by parents and others who are highly familiar with the child. Clearly, more research is needed on this important topic before these conclusions can be accepted, even though many examples consistent with the claim that interview bias has large effects on children's reports can be found in legal case files. This can be seen explicitly in the Kelly Michaels case, which, as of this writing, is being appealed on the basis of the defendant's contention that most of the children were subjected to relentless and single-minded interviews that were suggestive and even threatening. The following interview of one of the alleged victims, Child 8C, was conducted jointly by Mr. Fonolleras, a social worker, and by Detective Mastrangelo of the local police department:

Fonolleras: Don't be so unfriendly. I thought we were buddies last time.

8C: Nope, not any more.

Fonolleras: We have gotten a lot of other kids to help us since I last saw you. . . . Did we tell you that Kelly is in jail?

8C: Yes. My mother already told me.

Fonolleras: Did I tell you that this is the guy (pointing to Mastrangelo) that arrested her? . . . Well, we can get out of here real

quick if you just tell me what you told me the last time, when we met.

8C: I forgot.

Fonolleras: No you didn't. I know you didn't.

8C: I did! I did!

Fonolleras: I thought we were friends last time.

8C: I'm not your friend any more!

Fonolleras: How come?

8C: Because I hate you!

Fonolleras: You have no reason to hate me. We were buddies when you left.

8C: I hate you now!

Fonolleras: Oh, you do not, you secretly like me, I can tell.

8C: I hate you.

Fonolleras: Oh, come on. We talked to a few more of your buddies. And everyone told me about the nap room, and the bathroom stuff, and the music room stuff, and the choir stuff, and the peanut butter stuff, and everything. . . . All your buddies [talked]. . . . Come on, do you want to help us out? Do you want to keep her in jail? I'll let you hear your voice and play with the tape recorder; I need your help again. Come on. . . . Real quick, will you just tell me what happened with the wooden spoon? Let's go.

8C: I forgot.

Mastrangelo: Now listen, you have to behave.

Fonolleras: Do you want me to tell him to behave? Are you going to be a good boy, huh? While you are here, did he [Det. Mastrangelo] show you his badge and his handcuffs? . . . Back to what happened to you with the wooden spoon. If you don't remember words, maybe you can show me [with anatomical dolls present].

8C: I forgot what happened, too.

Fonolleras: You remember. You told your mommy about everything about the music room and the nap room, and all that stuff. You want to help her stay in jail, don't you? So she doesn't bother you anymore and so she doesn't tell you any more scary stories. (6/27/85 Appellate Court Brief)

This interview was characteristic of many of the state's interviews in the Kelly Michaels's case, with highly suggestive use of props and a relentless pursuit of only one hypothesis, often accompanied by bribes for disclosures and implied threats in the face of nondisclosure. Similar patterns of threats, bribes, and insinuations that their friends had already told investigators of the defendant's abusive behavior can also be seen in other cases (see Benedek & Schetky, 1987). Finally, these sorts of threats and bribes are not unique to prosecution interviews; similar examples can be found in defense interviews. For instance, in the Country Walk case, the following interview occurred between Samek, Frank Furster's attorney, and a 6-year-old child who had made multiple allegations:

Samek: You have been saying a lot of things about Frank and Iliana, haven't you?

Child: Yes.

Samek: I'm Frank's friend, and I want to help Frank, and I think you're lying. I think you're lying.

Child: No.

Samek: I don't think any of the things you are saying about Frank are true. Do you know what a lie is?

Child: When you—

Samek: No, look at me! You know what a lie is. What's a lie?

Child: When you say something that's not true.

Samek: OK, that's right. That's exactly what a lie is. I think you've been lying to me about Frank and Iliana. I don't think Frank and Iliana ever did anything to you. Frank didn't do anything to you, did he?

Child: Yes he did.

Samek: Frank never put his mouth on your penis, *did he?* (Hollingsworth, 1986, p. 76)

Clearly, the impact of that style of questioning, of being forced to look into the eyes of an angry and accusatory attorney in a strange and threatening courtroom, would seem threatening to a child. Yet, the child maintained his story that Frank Furster did have oral sex with him, a story that was later supported by Frank Furster's wife, Iliana. Thus, young children are apparently capable of accurately reporting what they witnessed at least some of the time, assuming that Iliana Furster's supporting testimony was itself truthful in this case.[5]

Anatomical dolls. Anatomical dolls are frequently used by professionals, including child therapists, police, child protection workers, and attorneys, when interviewing children about suspected sexual abuse. According to recent surveys, 90% of field professionals use anatomical dolls in their investigative interviews with children suspected of being sexually abused (Boat & Everson, 1988; Conte, Sorenson, Fogarty, & Rosa, 1991). Although we could find no national figures, it appears that expert testimony is often based on observations of children's interactions with these dolls (Mason, 1991). We discuss anatomical dolls in this section on interviews because issues regarding the degree to which dolls are suggestive have been raised by a number of commentators (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 critical event, thereby cuing recall and overcoming language and memory problems. A second rationale for the use of these dolls is that they are thought to overcome motivational problems of embarrassment and shyness. Children may feel more comfortable enacting an abusive event using the dolls than verbally recounting it. The dolls have also been used as projective tests. Some professionals claim that if a child actively avoids these dolls, shows distress if the dolls are undressed, or shows unusual preoccupation with the dolls' genitalia, this is consistent with the hypothesis that the child has been abused.

The use of these dolls, however, has raised skepticism among researchers and professionals alike. There are two related arguments frequently made against their use. The first is that the dolls are suggestive; they encourage the child to engage in sexual play even if the child has not been sexually abused (e.g., Gardner, 1989; Terr, 1988). For instance, a child may insert a finger into a doll's genitalia simply because of its novelty or "affordance." Another criticism is that it is impossible to make any firm judgments about children's abuse on the basis of their doll play because there are no normative data on nonabused children's doll play and no standardized procedures for their use (e.g., at which point in the interview they are introduced, whether they are introduced with their clothes on or off).

[5] Since we wrote this section, we have learned of arguments that Ileana Furster's "confession" might not have reflected her true opinion. She was held in solitary confinement for 1 year following her arrest. During this time, she consistently denied any allegations of abuse. Then, for a period of 2 months, she was visited on a daily basis by a friend of her lawyer, a priest, and a therapist. The latter is claimed to have made every effort to persuade her to turn state's evidence to save herself from a much more severe sentence (Nathan, 1993).

Because of these concerns, the use of these dolls for the purpose of providing legal evidence has been banned in a few jurisdictions until scientific data can be produced to attest to their validity. That research is beginning to be conducted. Since 1985, five studies have examined the degree to which sexually abused children's interactions with the dolls differ from those of nonabused children. The findings of these studies are inconsistent.

August and Forman (1989) observed the spontaneous doll play of 5- to 8-year-old girls who were suspected of being abused or who were not abused; they used the dolls to retell a story. Two raters who were aware of the children's status conducted ratings of the children's behavior. Abused children showed more avoidance of the dolls when asked to tell a story, and they engaged in more sexual activities than did the nonabused children.

White, Strom, Santili, and Halpin (1986) conducted interviews using anatomical dolls with 2- to 6-year-old abused and nonabused children. Raters, unaware of the status of the children, were more likely to rate the abused children as showing abuse. The two groups of children also differed in the quality of doll play; sexually abused children showed excessive interest in the anatomical parts and in their demonstration of sexual acts. Nonabused children showed no unusual behavior in relation to sexual play with the dolls.

However, using a similar methodology, Realmuto, Jensen, and Wescoe (1990) reported that raters could not reliably distinguish between abused and nonabused children's play with anatomical dolls. Similarly, Cohn (1991) compared the doll play behaviors of children referred for assessment of sexual abuse with a nonabused group of children. All subjects were aged 2–6 years. The two groups did not differ on measures of frequency of sexually explicit behaviors. For example, 11% of the referred children and 17% of the nonabused children inserted their fingers into the dolls' private parts. Finally, although Jampole and Weber (1987) found that 90% of their abused sample engaged the dolls in sexual activity more than did a nonabused sample, these sexually explicit behaviors were also observed in 20% of the nonabused children.

The divergent findings of these studies may reflect two factors. First, in most studies, interviewers were aware of the status of the children, a condition that could have influenced their subsequent interactions with the children, especially when playing with the dolls (Wolfner, Faust, & Dawes, 1993). Second, in most studies, children "suspected" of being abused are compared with children "not referred" to sexual abuse clinics. Because there is rarely any validation of these diagnostic categories, it is likely that some of the children are misclassified in terms of group membership.

A second set of studies have examined in greater detail how children who are not suspected of being abused play with anatomical dolls. Sivan, Schor, Koeppl, and Noble (1988) observed a middle-class sample of 3- to 8-year-olds interact with anatomical dolls. Role playing with explicit sexual activity was not observed. Glaser and Collins (1989) conducted a similar study on middle-class children (aged 2–6 years). Five percent of the children refused to play with the dolls, and 35% showed some reticence or avoidance of the dolls. Five percent showed explicit sexual play. On further investigation, 3 of these 5 children had been exposed to either pornographic literature or had observed sexual activity. There were no apparent explanations for the interactions of the other 2 children. Thus, premature exposure to sexuality rather than sexual abuse could account for some children's explicitly sexual interaction with anatomical dolls.

A third study, conducted by Everson and Boat (1990), examined the interactions of a socially diverse sample of children. Their focus was on the degree to which 2- to 6-year-olds used the dolls to show suggestive or explicit sexual behavior when they played with anatomical dolls in the presence of an interviewer versus when they were alone. Although none of the 2-year-olds showed suggestive or clear intercourse positioning, this did occur for 9% of the 4-year-olds and 16% of the 5-year-olds. When the data were analyzed in terms of race and socioeconomic status (SES) of the child, only low-SES Black boys showed clear intercourse positioning in the interviewer-present condition.

In order to determine whether their subjects' initial exposure to the anatomical dolls had any long-lasting influences, Boat, Everson, and Holland (1990) interviewed a subsample of mothers of the 3- to 5-year-olds who had played with the dolls 2 weeks previously in the Everson and Boat (1990) study just described. None of the mothers of 5-year-olds reported any noticeable behaviors that might have been related to the doll play. However, 37% of mothers of 3- and 4-year-olds reported that there were behavioral sequelae to the doll play. This was mainly demonstrated in an increase in sexual interest that involved talking or asking about sexual parts. Furthermore, 50% of the mothers of 3- and 4-year-old children believed that their children's behaviors were more sexually focused. However, these were considered to be benign behaviors that would not lead to a later interpretation or question of sexual abuse. Nevertheless, these data do indicate that after one exposure to the dolls, preschool children's behaviors were noticeably different in the eyes of their mothers.

To our knowledge, there has been only one study in which dolls were used to probe children's memories for a neutral event that involved interacting with a male confederate (Goodman & Aman, 1990). Three- and five-year-olds were questioned 1 week after this interaction. Children were questioned under one of four conditions: anatomical dolls, regular dolls, regular dolls that the child could not touch, or no dolls. The children were encouraged to use the dolls to show what had happened. Recall of events and answers to objective and misleading questions were similar across all conditions. Although this study showed that children in the anatomical doll condition did not report more sexually related events, it is important to note that the dolls did not facilitate accurate memory recall of this neutral event, indicating that their mnemonic value may be limited.

To summarize, the data on anatomical dolls are equivocal. Some studies have shown clear differences between abused and nonabused children's interactions with the dolls. Some researchers claim that nonabused children rarely if ever show sexually explicit play with the dolls, whereas others argue that a small proportion do show such behaviors. Although these rates are low for middle-class samples, they increase in more socially diverse samples. Our reading of the literature suggests that the techniques for using anatomical dolls have not been developed

to the level that they allow for a clear differentiation between abused and nonabused children. It seems that for a small number of nonabused children, the dolls are suggestive in that these children engage them in sexual play.

It is not clear why these studies have yielded such divergent findings, although they do differ in the age groups studied, the procedures used, and the demographic characteristics of the samples. This divergence points to the need for additional research as well as to the need for explicit procedures to govern the use of the dolls by interviewers. Until such time that research is available, the dolls ought to be used with great caution.

Caveat lector. In the debate over the suggestibility of dolls, one problem has been overlooked: One cannot generalize from studies of anatomical dolls to actual sexual abuse interviews because the contexts for the presentation of the dolls is much different in research settings than in forensic and clinical settings. Transcripts of therapy sessions with children suspected of being sexually abused reveal the following practices: naming the dolls after defendants, berating the dolls for alleged abuses against the children, assuming the role of fantasy characters in doll play, and creating a persistent atmosphere of accusation. In the experimental studies of anatomical dolls, nonabused children were not 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, the children who were alleged to have been abused were exposed to the dolls repeatedly prior to coming to the research setting, often amid a stream of suggestions from parents and interviewers about various sexual themes. That they played with the dolls differently from nonabused children who lacked this prior experience could have been the result of the prior experience rather than anything inherent in the way an abused child would play with the dolls for the first time. Thus, the literature on anatomical dolls does not reveal whether nonabused children would interact with the dolls differently from abused children if the former were subjected to the same preexperimental experience of the abused children (i.e., multiple interviews with the dolls in the context of discussing sexual matters). This raises the possibility that a child behaved sexually with the dolls, not because he or she was abused but because of prior sexual discussions in conjunction with previous doll use—a possibility independently raised by Wolfner et al. (1993).

Summary of interviewing studies. The studies on interviewing provide evidence that suggestibility effects are influenced by the dynamics of the interview itself, the knowledge or beliefs possessed by the interviewer (especially one who is unfamiliar with the child), the emotional tone of the questioning, and the props used. Children attempt to be good conversational partners by complying with what they perceive to be the belief of their questioner. Their perceptions, and thus their suggestibility, may be influenced by subtle aspects of the interview such as the repetition of yes–no questions, but their compliance is evidenced most fully in naturalistic interview situations in which the interviewer is allowed to question the child freely; this gives the child the evidence to make the necessary attributions about the purposes of the interview and about the intents and beliefs of the interviewer.

Observations of interactions in the legal arena highlight the fact that children who testify in court are not interviewed in sterile conditions such as those found in many of the experiments we have reviewed. They are usually questioned repeatedly within and across sessions, sometimes about an ambiguous event by a variety of interviewers, each with their own agenda and beliefs. Children are sometimes interviewed formally and informally for many months preceding an official law-enforcement interview with anatomical dolls, providing an opportunity for the child to acquire scripted and stereotypical knowledge about what might have occurred.

Social and Cognitive Mechanisms in Lying

An equally important consideration in evaluating the suggestibility of child witnesses concerns the conditions under which children *consciously* and deliberately distort the truth because of a variety of social and motivational factors that extend beyond the interview. Historically, it was felt that lying[6] was beyond a young child's cognitive capability because it required a greater degree of decentration than preschoolers exhibited (e.g., Piaget, 1926).

Since the time of Piaget, much progress has been made in understanding the development and definitional features of deception. It is beyond the scope of this review to chronicle this progress because much of it is not germane to the main theme here: young children's proneness to suggestion in response to powerful motives.[7]

With advances in the understanding of young children's cognitive sophistication, there is now evidence that even very young children sometimes do lie, with full appreciation of the differing perspectives of their listeners. For example, 88% of 3-year-olds who were instructed not to peek at a toy proceeded to peek. When asked if they had peeked, only 38% admitted to it, prompting the investigators to conclude that "thus, we have

[6] Here, we use the term *lying* to refer to the deliberate, conscious production of a response that the child believes to be incorrect for the purpose of achieving a goal, namely, misleading the listener to believe it is correct. We do not assume any malintent on the part of the "liar" that the term sometimes connotes. Also, we do not assume when we use the term *lie* that the child has solved the philosophical problem of inferring the contents of the listener's mind (Chandler, 1989). Philosophers like Flanagan (1992) have assumed that to engage in a minimal lie, the perpetrator who believes x must do something or omit something with the intent of making the listener think *not* x: "A minimal lie requires understanding the complex relation between actions (pointing in the wrong direction, speaking falsely), and the production of false beliefs in one's audience" (Flanagan, 1992, p. 15). We merely use the term *lying* for ease of reference in discussing research that does not meet this high standard of cognitive awareness.

[7] For recent discussions of the definitional complexity of deception, see reviews and analyses by Leekam (1992), Perner (1991), and Sodian (1991). For an analysis of the link between emergent theories of mind and children's understanding of the beliefs of those they try to deceive, see Mitchell and Thompson (1986), Chandler, Fritz, and Hala (1989), Perner (1991), and the recently published book by Whiten (1991).

some evidence . . . that deception strategies are adopted at early ages" (Lewis, Stranger, & Sullivan, 1989, p. 442). Although some researchers have claimed that higher order deception (the child infers the state of mind of the people they are trying to deceive and tries to inculcate a false belief in them) does not appear prior to the elementary school years (see Perner, 1991), others have shown that most 4-year-olds have some degree of cognitive sophistication in attempting to deceive, whereas 3-year-olds do not (Leekam, 1992).

We now focus on studies that have examined preschoolers' deception, ignoring whether behaviors are more appropriately construed as "sabotage," "deceit," "tricking," "politeness," or "tact." Furthermore, we avoid delving into distinctions that have occupied "theory of mind" scholars, such as lying versus telling a lie and minimal lies versus deception.

Recent research has sought to examine the specific conditions that may foster lying. Five motivations to lie or tell the truth have been studied: avoiding punishment; sustaining a game; keeping a promise (e.g., to protect a loved one); achieving personal gains (e.g., rewards, being accepted in a group); and avoiding embarrassment. Existing data show that not all motivations produce comparable levels of lying and truth telling.

Lying and truth telling to avoid punishment. Children will lie about events when the operative motives are sufficiently salient, and they will do so at an earlier age than previously assumed to be the case. Mothers report that the most frequent motivation for their 4-year-olds to lie is to avoid punishment (Stouthamer-Loeber, 1987), a finding echoed in the recent findings of Bussey (1992).

Lying and truth telling to sustain a game. Some children can be induced to tell a lie in the context of a game. For example, an adult experimenter pretended to find a watch left behind by the teacher (Ceci, DeSimone, Putnick, Lee, & Toglia, 1990). After showing the child the watch, the child was told they were going to play a game of hiding it from the teacher. The child was told the game was a secret and was instructed not talk to anybody about it. Later, the returning teacher asked the child who had taken her watch. Only 10% of the preschoolers lied to sustain this game. Tate and Warren-Leubecker (1990) and Pipe and Goodman (in press) have reported similar figures. However, when the motivational salience of the experimental procedure was increased by having a well-known adult coach the child to tell a lie about playing with a toy, 35% of 2- to 8-year-olds lied to sustain a secret game (Tate, Warren, & Hess, 1992). It appears that the degree to which children will lie to sustain a game is context dependent and that the use of stronger coaching will result in higher rates of deception.

Keeping promises. There is consistent evidence that children as young as 3 years of age will omit important information about transgressions and accidents if adults ask them to do so (see Pipe & Goodman, in press, for a recent review). For example, in one study an adult spilled ink on a pair of gloves the child was wearing and told the child that she (the adult) would "get into trouble" if anyone found out. Subsequently, 42% of the 5-year-olds claimed not to know who spilled the ink, and 25% maintained ignorance on repeated questioning 10 days and 2 months later (Wilson & Pipe, 1989). Peters (1991c) reported similar results. Four- to ten-year-olds witnessed a staged event of a stranger who stole a book and were asked to keep the theft a secret. When the children were asked by the owner of the book whether they had seen who took it, 82% either delayed reporting the theft or never reported it. The most common reason given by the children for not disclosing was to honor the stranger's secret and to avoid getting him into trouble. Finally, some of the children in the Clarke-Stewart et al. (1989) study were told by Chester that he would lose his job if his boss learned that he had played with the dolls. Sixty-nine percent kept the secret when they were interviewed by a neutral interviewer. However, they all eventually revealed the secret when asked suggestive questions.

If children will lie to protect a stranger, they should do so even more readily to protect a loved one. Results of one study support this hypothesis (Bottoms, Goodman, Schwartz-Kenney, Sachsenmaier, & Thomas, 1990). When mothers of 3- and 5-year-olds broke a Barbie doll, only 1 of the 49 children mentioned this to an interviewer who asked what happened. Furthermore, when asked specific questions about the event, 5-year-olds did not tell the secret, even when asked leading questions.

Lying and truth telling for personal gain. Sometimes children will lie to gain a material reward or to gain acceptance in a group. For example, in the Pettit et al. (1990) study, 7 children were absent from school on the day of the staged event but were in the same classrooms as the children who saw the event. Only 1 of the 7 children said that he was not present on the day the staged event occurred. Three of the 7 children gave accounts to the interviewers that were indistinguishable from those of children who had seen the event. These 7 children probably overheard discussions of the event and wanted to be part of the group.

Material and psychological rewards do not need to be of a large magnitude to be effective. Children as young as 3 years of age will engage in sabotage behaviors to gain some reward (e.g., covering up a treasure that is in jeopardy of being discovered by a puppet), but they will not engage in verbal deceit for another year (Leekam, 1992). In a study by Ceci and colleagues, more than 50% of nursery school subjects lied to obtain a gumball as a prize by falsely claiming that they had won a game while the interviewer was out of the room (Ceci, Leichtman, Putnick, & Nightingale, 1993). Control children who had played the game but were not offered any prizes accurately reported that they had not won the game, so simple memory failure can be ruled as an explanation of children's erroneous claims.

Lying and truth telling to avoid embarrassment. Not all behavior is regulated by external outcomes, rewards, and fears of reprisals. Self-regulatory mechanisms also effectively direct behaviors (Bandura, 1986). In the context of lies, the most relevant self-regulators are guilt, shame, and pride. Although children prior to the age of 7 (Bussey, 1992; Leekam, 1992) appear to be inferior to older children at inferring some of these self-regulatory states (e.g., appreciating that a sense of pride results from telling the truth), even 4-year-olds distinguish between statements meant to minimize the embarrassment of another (e.g., "I like your new hairdo") and those meant to cause distress (Bussey, 1992).

In an effort to study lying to avoid personal embarrassment,

Ceci, Leichtman, Putnick, and Nightingale (1993) instructed two parents to kiss their 3-year-olds while bathing them the evening before being interviewed. During an interview in which their parents were absent, the children were told that it was naughty to let someone kiss them when they did not have any clothes on. They were then asked, "No one ever kissed you when you didn't have any clothes on, did they? . . . No one kissed you last night when you were in the bathtub, did they?" (Ceci et al., 1993). These instructions provided a motivation to make "errors of omission," or withhold information about an event portrayed as taboo, in order to avoid the embarrassment of having done something naughty. Immediately following the child's reply, he or she was told it was okay to be kissed by a parent or someone they knew. Later, the children were asked by their parents whether they had been kissed while being bathed. In a different condition, two children who had *not* been kissed during their baths were told at the start of the interview that parents who love their children often kiss and hug them while they are in the tub and asked them the following: "Your mommy kissed you when she bathed you last night, didn't she?" (Ceci et al., 1993). Later, their parents also asked this question. The purpose of this condition was to provide a motivation to make "errors of commission" in order to avoid embarrassment.

Initially, both children who were told that it was naughty to allow an adult to kiss them while being bathed replied that they had *not* been kissed. Later, when a parent interviewed them alone and asked if they had been kissed while being bathed, they affirmed that they had, offering specific and accurate details (e.g., "Yes, I think mommy kissed me three times in the tub last night"). Interestingly, the children quickly added a codicil that was nearly a verbatim restatement of the interviewer's assurance: "But it's OK because I know her" (Ceci et al., 1993). Of the 2 children who had not been kissed during the evening bath, 1 child reported that she had been but reversed her report when interviewed by a parent alone. The results of this case study indicate that occasionally, children will consciously distort the truth about events that were allegedly perpetrated to their bodies. Both errors of omission and errors of commission were produced by the strong motives used by these researchers.

It is imprudent to make too much of this research because of its limited sample size and scope. Yet, the fact that 1 out of 2 children made an error of commission about an abuse-relevant action raises doubts about claims that fewer than 1% of children can be led to report false touching (Goodman et al., 1990). Until now, researchers who have claimed that children cannot be coached to distort their testimony appear to have tilted the odds toward finding truthfulness among preschoolers by implicitly using motives that favor a truthful outcome (e.g., Goodman et al., 1990; Saywitz et al., 1991). There were no motives for the child to make false disclosures in these earlier studies. It might even be claimed that in such studies, there are implicit motives to correctly report because to do otherwise would bring embarrassment. If children were to distort what they had witnessed and claim to have been sexually touched when they were not, this could be expected to result in embarrassment, thus tilting the motivational structure toward truthful reporting. Contrast this approach with a child asked to make errors of commission to avoid embarrassment (e.g., "He kissed you because he loves you, didn't he?") or to an approach in which a child is asked to make errors of omission to avoid embarrassment (e.g., "No one ever touched you there, did they?"). Therefore, it is important in making sense of the disparate findings to be mindful of the operative motives, both implicit and explicit.

In sum, the most recent research on lying has attempted to approximate real-life crime contexts by weaving affect and motive into studies of recollection and by using highly familiar contexts such as observing loved ones break toys or being kissed while in the bathtub. Young children will consciously distort their reports of what they witnessed, and they will do so more in response to some motives (e.g., fear of reprisal and avoidance of embarrassment) than others (e.g., to sustain a game, gain rewards).

Earlier research has shown that children do not have the cognitive skills to engage in deception at early ages (see chapters in Ceci, Leichtman, & Putnick, 1992). However, these earlier studies used paradigms that required a high level of cognitive skill to carry out the ruse. For example, children were asked to pretend that they liked something that they did not in fact like (Ceci, Leichtman, & Putnick, 1992). Because of the sophisticated cognitive skills needed to engage in pretense play, such paradigms seem to have underestimated preschoolers' ability to use deception. Alternatively, it may be that even when young children have the requisite cognitive skills to engage in pretense play of the sort required by these paradigms, they will refuse to do so because it would conflict with their negative affect (i.e., claiming to like something that they do not like), something they are unwilling to admit at a young age for reasons having more to do with social than cognitive skills.

Primacy of Cognitive Versus Social Mechanisms

Our review of the literature indicates that there are a variety of cognitive as well as social and motivational mechanisms that influence children's suggestibility. The issue that we address in this section concerns which set of factors may best account for children's suggestibility and for age-related differences in children's suggestibility and how these factors may interact.[8]

There have been few studies that have directly contrasted the relative influence of social and cognitive factors on children's suggestibility. Ceci et al. (1987) showed that when an erroneous postevent suggestion was supplied by an adult, it resulted in significantly more incorrect recognitions of the original event than when it was supplied by another child, confirming the importance of social factors, such as interviewer prestige in the suggestibility effect. However, even when a child supplied the erroneous suggestion, it still resulted in more incorrect recognitions than was found for children in the nonmisled group. These data suggest that although social factors are clearly important, they do not fully account for suggestibility effects. Young children's memory is affected by erroneous suggestions over and above the problems created by social factors. Of

[8] As we show, it is somewhat misleading to speak of social and cognitive factors as though they operate independently in producing suggestibility effects.

course, it is possible that even a child interviewer who provides erroneous suggestions presents some degree of social influence.

The degree to which social and cognitive factors influence suggestibility may be gleaned from a considerations of the methodology for assessing suggestibility. A review of the studies described in this article makes clear that there is much variability in the manner in which misinformation is provided and in the methods used to assess suggestibility. In some studies, misinformation was incorporated into the interview questions that occur after the subjects have viewed the critical event (e.g., Warren et al., 1991). In other investigations, misinformation was presented prior to testing (e.g., Ceci et al., 1987; Zaragoza, 1991; Zaragoza et al., 1992). In addition, the timing of the misinformation varied across studies: Misinformation could occur minutes to several hours or even days relative to the target event. Although all ways of assessing suggestibility may be related, they need not be. For example, a child who acquiesces to a leading question that is presented for the first time during the testing session may at a later time not exhibit suggestibility (e.g., Cohen & Harnick, 1980). This would be an example of a social factor affecting the report (e.g., going along with an interviewer's suggestion because of a perceived pressure to conform) rather than a cognitive factor (memory alteration). Thus, acquiescence to a leading question provided at the time of testing does not in itself imply that the misinformation contained in the leading question has been incorporated into the memory. Similarly, it is possible that a child who initially resisted misinformation during testing might on subsequent occasions show evidence of suggestibility. Memory impairment may require a delay interval in order for sufficient forgetting to occur before the erroneous suggestion can alter the original memory (Belli, Windschitl, McCarthy, & Winfrey, 1992). Thus, presenting information shortly after the original information may result in a different mechanism than when misinformation is presented much later.

Focusing on cognitive versus social factors in the suggestibility effect obscures the possibility that both factors interact in producing it. For instance, it is possible that the degree to which social factors play a role has a cognitive basis. When memory traces are weak (or when there is no memory for the original event), children may be more compliant and willing to accept suggestions because there is no competing trace to challenge the suggestion. On the other hand, when the traces are strong, the child (or adult) is less likely to incorporate misleading suggestions into memory. Thus, Warren et al. (1991) reported that lower recall was significantly related to greater suggestibility. Their subjects were most easily misled about the details that were omitted in their original recalls. Furthermore, there was a significant correlation between subjects' free recall of the event and the degree to which they changed their answers to their questions on the second round of questioning, a reflection of how cognitive factors affect social manipulations (i.e., children are most susceptible to suggestible influences when their memories for events are weakest).

In addition to cognitive factors underpinning the effectiveness of social factors, it is also possible that social factors underpin the effectiveness of cognitive mechanisms in producing suggestibility (e.g., a child may attend more to suggestions from authority figures), thus ensuring greater encoding. However, this is a hypothesis in need of data.

Finally, it is possible that a child's report may initially be the result of some social factor, but over time the report may become a part of the actual memory. Earlier, we described the "Sam Stone" study in which preschool children were given stereotypical knowledge about a clumsy character (Ceci, Leichtman, & White, in press). Children later used this knowledge to reconstruct what Sam Stone *might* have done. On repeated postevent questioning, however, the children often became more convinced that the clumsy events had actually occurred, as opposed to *might* have occurred. Over 70% of 3- and 4-year-olds incorporated some of these postevent suggestions into their reports, and nearly 20% of them refused to relinquish their claims when the interviewer tried to talk them out of it. In the legal arena, in response to strongly suggestive—even pressurized—interviews, children may initially realize that they are providing the interviewer with an erroneous account, but after repeated retellings to different interviewers, the erroneous account may become so deeply embedded as to be indistinguishable from an actual memory. This is precisely the point of contention between those who maintain that genuine memory impairment can result from persistent erroneous postevent suggestions (e.g., Ceci et al., 1987) and those who argue that the original memory of the event is unimpaired (e.g., McCloskey & Zaragoza, 1985a, 1985b; Zaragoza, 1991).

Although cognitive and social factors may both play a role in suggestibility effects, the important question is, When do social variables become important developmentally and to what degree do social or cognitive variables account for suggestibility effects? Of particular interest to our discussion is whether age-related changes in the relative importance of these social and cognitive factors can account for the situations when children show heightened levels of suggestibility. Clearly, much more research is needed to gain a fuller understanding of the boundary conditions.

Finally, it needs to be stressed that regardless of the outcome of the debate over cognitive versus social mechanisms, both camps agree that there are pronounced age differences in the accuracy of children's reporting in the face of misleading questions. The disagreement is confined to the causal mechanisms that underlie the observed suggestibility effects, not to whether suggestibility effects exist, a position that almost all endorse.

Causal Mechanisms: Biological Factors

No review of suggestibility would be complete without inquiring into the biological variables that have been posited to account for age differences in suggestibility. Here, we focus on the factor of arousal or stress. This issue is of importance in our discussion because children who provide testimony often recall events that occurred under highly stressful conditions, and often these children become stressed as a result of the interviewing contexts. Thus, it is useful to examine the degree to which children provide accurate reports under these conditions.

Stress, Memory, and Suggestibility

There is a growing but controversial literature on the relation between arousal and children's memory, and some of these studies have included tests of the relation between arousal and suggestibility. Some researchers claim that high levels of arousal are beneficial for children's testimonial accuracy as well as for their resistance to suggestions (e.g., Goodman, 1991); others maintain that high levels of stress are debilitating, resulting not only in less memory but also greater suggestibility (e.g., Peters, 1991a, 1991b, 1991c); and some researchers argue that stress and arousal are unrelated to memory or suggestibility effects (e.g., Oates & Shrimpton, 1991; Steward, 1989).

The position that stress facilitates recall is consistent with the phenomenon known as "flashbulb memories" (e.g., R. Brown & Kulik, 1977; Linton, 1982). The classic example of this is the claim that most people can remember where they were when they heard of President Kennedy's assassination. Events that are emotionally arousing are thought to receive privileged encoding; high emotional salience is associated with a "print now" mechanism that reputedly permits the core details of affectively valenced events to be automatically encoded (R. Brown & Kulik, 1977). A physiological explanation for flashbulb memories posits that shock releases hormones such as adrenaline that drive up the plasma glucose level, which may be the basis for the enduring memory (e.g., Gold, 1987).

There is some support for the hypothesis that high levels of emotional arousal are associated with accurate and consistent recall. High arousal among college subjects in the aftermath of the space shuttle *Challenger's* explosion was associated with greater consistency of reporting over a 3-year interval (Bohannon, 1988). There are some child studies that are consistent with concept of flashbulb memories. Terr (1983, 1990) has described numerous clinical cases in which individuals exposed to traumatic experiences in early childhood, such as a sibling's murder or their own rape, were able to recollect their experiences in detail. Children's reports of the *Challenger* disaster over a 2-year period were more consistent among those who reported higher emotional responses to the disaster than those who reported lower emotional responses (Warren & Swartwood, in press).

In line with these results, Goodman and her colleagues conducted four experiments to examine children's memory for stressful events involving inoculation and venopuncture (Goodman, Hirschman, Hepps, & Rudy, 1991). In two studies, stress was beneficial to recollections of children, and in one of the two, high stress was associated with resistance to suggestion. In the first of these studies, 3- to 6-year-olds received an inoculation; their emotional state during this procedure was categorized on a 6-point scale ranging from *very relaxed* to *very frightened*. The children's recollections of the inoculation were tested 3–9 days following the shot. The highest levels of arousal were associated with the most accurate recall and the lowest levels of suggestibility. In the second study, children who did not receive an inoculation but who went to the clinic to have a decal rubbed onto their arm and leg were also interviewed 3–9 days after this procedure. Their responses were compared with those of the inoculated children just described. Of the eight major analyses performed, one resulted in a significant difference between inoculated and noninoculated children: Noninoculated children (i.e., those with lower levels of stress) recalled more incorrect information than did the inoculated children.

There is another set of studies in which stress was associated either with no differences in subsequent memory or with detrimental effects on memories and their resistance to suggestion (Kramer, Buckhout, Fox, Widman, & Tusche, 1991). This literature has fueled the growing discontent over the flashbulb metaphor because of the presumed uniqueness of its mechanisms and the consequences related to completeness of report, accuracy, and immutability (see Bohannon, 1988; McCloskey, Wible, & Cohen, 1988; Pillemer, 1990; Winograd & Killinger, 1983).

Several studies in the child literature have also failed to show any association between stress and memory. Even though there was a relation between children's emotional response to the *Challenger* disaster and the consistency of their recalls over a 2-year period, there was no relation between emotional response and number of core details reported after 2 years (Warren & Swartwood, in press). Steward (1989) reported that in initial interviews, children who received a painful medical procedure disclosed significantly more information about being touched than did children who did not receive this procedure. However, 6 months later, there were no differences in the reports of the high-pain group and the group that did not receive this procedure (see Goodman, Hirschman, et al., 1991, Studies 1 and 4, for similar findings).

In some cases, stress is related to impoverished recall on some but not all measures. Vandermaas (1991) found that high levels of stress associated with dental procedures had no effect on the accuracy of 4- to 7-year-olds' responses to specific questions, although it did have a detrimental effect in terms of their free recall; it led to diminished recall among the youngest children. Ornstein et al. (1992) correlated children's distress (as rated by nurses, doctors, and parents) to children's immediate and delayed recall of their visit to their pediatrician. Of the nine possible correlations, only one was significant. Parental stress ratings of the older children were negatively correlated with delayed recall; the more stressed children recalled fewer details of their visit than did less stressed children.

Several studies have shown a more consistent association between high levels of stress and low levels of recall. Bugental, Blue, Cortez, Fleck, and Rodriguez (1992) have provided a direct test of the hypothesis that stress influences encoding by recording 5- to 10-year-olds' electrodermal responses and heart rate changes as they watched a mildly stressful video showing frightened faces. Details of the video that followed heightened arousal were remembered less well than details presented prior to the elevated arousal. Thus, if it is argued that higher levels of stress result in better encoding, the evidence does not seem to support this (Oates & Shrimpton, 1991; Warren & Swartwood, in press).

The most extensive evidence in favor of the hypothesis that stress impairs children's memory and increases their suggestibility has been provided by Peters (1991a, 1991b), who conducted five experiments of children's recollections of different stressful events. After testing close to 400 children in which