

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 OGP000 A1181 to Maggie Bruck, and by National Institutes of Child Health and Human Development Grant RO1 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, bed-wetting, 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. Interwoven 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 (sex-

ual 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 & Hammond, 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. Münsterberg (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. Burt 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 Burt’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 suggesti-

bility 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, Togliola, & 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-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 procedure was mildly suggestive, breaking down their shaky differentiations, and the 4-year-olds had more fragile boundaries than did the 6-year-olds. In both case studies (summarized at the beginning of this article), children's disclosures became increasingly bizarre and incredible. This evolution could have been caused by the interviewers not drawing the children back to reality when they made "fantastic" claims and perhaps as a result, their allegations moved from fantasy to reality for the children who came to believe themselves.

Another area of research focuses in greater 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; Lindsay 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 relax 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 enjoiner that "they will feel better once they tell," then repeated and stronger enjoiners (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 Iliana 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, Koepl, 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 lying⁶ 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 be-

cause 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 consideration 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 post-event 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

Table 1
Summary of Stress Studies

Study	Memory effect	Suggestibility effect
Goodman, Hirschman, et al. (1991, Exp 2)	+	0
Goodman, Hirschman, et al. (1991, Exp 3)	+	+
Peters (1991a, Exp 1)	-	
Peters (1991a, Exp 2)	-	
Bugental et al. (1992)	-	
Peters (1991a, Exp 3)	-	
Peters (1991a, Exp 4)	-	
Peters (1991a, Exp 5)	-	-
Vandermaas (1991)	-	
Ornstein et al. (1992)	-	
Goodman, Hirschman, et al. (1991, Exp 1)	0	0
Goodman, Hirschman, et al. (1991, Exp 4)	0	0
Warren and Swartwood (in press)	0	
Oates and Shrimpton (1991)	0	0
Steward (1989)	0	

Note. To save space, only the first author's or the first and second authors' last names appear in the table citations. Exp = experiment. + = significant positive associations between stress and either memory or suggestibility. - = significant negative associations between stress and either memory or suggestibility. 0 = no significant associations between stress and either memory or suggestibility.

several measures of arousal were used, he concluded that high arousal levels can at times impair the eyewitness performance of children and that high arousal levels are never associated with increased recall. Experiment 5 of Peters (1991a) is most pertinent to this review because it included a suggestibility manipulation.

Five- to 10-year-olds were performing a card-sorting task when they unexpectedly heard either a fire alarm (high-stress group) or a loud radio (low-stress group). Measures of blood pressure and pulse rate taken before and after the onset of the noise showed a significant elevation only for the fire-alarm group. When children's memory was tested, no differences were found between the two groups on free recall, but the radio listeners gave more accurate responses to the objective questions than did the fire-alarm group. Recall was distorted by misleading questions in both groups, but the difference was significantly greater in the alarm condition. Thus, the stress of the alarm condition coupled with exposure to misleading questions reliably reduced accuracy of children's recall.

Unfortunately, few other studies have examined the effect of stress on suggestibility. Four studies showed no relation between these variables. Oates and Shrimpton (1991) found no differences in suggestibility between children who had received a blood test (high-stress group) or children who had interacted with a friendly stranger (low-stress group). Furthermore, there were no correlations between children's self-rated stress and their performance on suggestive questions or any other memory measures. Finally, although Goodman, Hirschman, et al. (1991) found that stress led to resistance of sugges-

tions, this was obtained only in one of the four studies reported. For the three other studies, all of which involved recall of events related to inoculations or to venopuncture, stress was unrelated to suggestibility.

Summary of Stress and Memory Studies

To summarize, 15 studies that examined the relation between stress and memory in children are reviewed in Table 1. In only two cases were there reports that high stress was beneficial to memory. In the remaining 13 studies, there was either no effect of stress on memory accuracy ($n = 5$) or high levels of stress were detrimental to memory ($n = 8$). Finally, only 6 of the 15 studies examined the effects of stress on suggestibility. Most ($n = 4$) studies reported no associations between stress and suggestibility. Two studies did show an association, but the direction of the correlation was different for each. It is difficult to resolve these inconsistent data, but we offer a few possible explanations.

First, there are clear differences among studies in the ways in which stress was measured and operationalized. Some studies relied on self-reports, whereas others relied on measures of physiological changes in arousal, such as salivary norepinephrine and immunoglobulin A levels. It is important to note that there were only moderate correlations between objective rating scales and various types of self-reports and that the correlations between different *subjective* reports of stress (parents', children's, and other observers') were often weak (Ornstein et al., 1992; Vandermaas, 1991). Thus, there were significant problems in terms of the validity of the classification of subjects according to stress. Even when these methods may be valid, however, there still is the problem of whether what is counted as a stressful event in one study is comparable to what is counted as a stressful event in another.

Interpretations of these results are also complicated because the accuracy and consistency of memory may be curvilinear across factors that are assumed to be correlated with levels of stress (Pillemer, Koff, Rhinehart, & Rierdan, 1987) as well as asymmetric across processing loci. Concerning the first claim, it may be that both the child who has received many prior inoculations and the child who has received no prior inoculations will differ from the child who has received a moderate number of prior inoculations. The former groups may have weaker emotional reactions to the inoculation experience, either because of an absence of prior experience that would result in little stress at the time of the first shot or so much experience receiving shots that the procedure has lost some of its fright. By contrast, the child who has experienced some, but not many, shots may be most stressed by the prospect of another shot. Pillemer et al. (1987) found that very strong as well as very weak emotional reactions may be associated with less subsequent recall among adults than moderate levels of stress (see also Defenbacher, 1991, for a recent review of the adult memory-affect relationship and its consistency with the Yerkes-Dodson law).

Concerning the claim that stress may exert asymmetric effects on memory at different processing loci, it is possible that stress at the time of encoding *could* aid the storage of information, whereas stress at the time of retrieval could impair access

to the contents of storage. For instance, Goodman, Levine, Melton, and Ogden (1991), in the amicus brief to the Supreme Court in *Craig v. Maryland*, noted that children's stress during testimony (i.e., during retrieval) could reduce willingness to articulate their memories. However, it is important to underscore the word *could* because the evidence that stress aids encoding is unconvincing.

One suggested resolution to the ongoing debate concerns a consideration of the type of information that is being erroneously suggested as well as the type of memory test used. Christianson and Loftus (1987) found that high arousal was associated with improved memory for central information in adults but that it was associated with a diminished ability to remember the specific details of the traumatic scenario. Similarly, Goodman et al. (1990) found that memory for central information is harder to bias through suggestive questioning. However, in the one study that examined suggestibility for central actions, high levels of arousal were also associated with heightened suggestibility. Oates and Shrimpton (1991) reported that children's suggestibility for actions or central information was greater than their suggestibility for the people with whom they interacted.

A compromise position may emerge on the basis of Pillemer's (1990) view that there is no mechanism inherently unique to the formation of flashbulb memories. Rather, the shock and salience of flashbulb experiences may result in their greater rehearsal and elaboration, which results in greater retention. A key determinant of whether arousal will be associated with subsequent age differences in memory and suggestibility may be moderated by how much the children know about the event. In two studies, children who had a fuller understanding of an emotionally arousing event remembered more than did children with less understanding of the event (Pillemer, in press; Warren & Swartwood, in press). According to this view, stressful experiences may be differentially rehearsed and elaborated by different age groups, and this could lead to different levels of memory and perhaps to differential suggestibility effects.

This interpretation may start to disambiguate some of the results previously reviewed. It is possible that stress levels are sometimes not associated with suggestibility effects or memory deficits because the stress manipulations are confounded with knowledge differences. In most studies, children who were classified as stressed usually received a procedure that was highly scripted and perhaps rehearsed with parents before and after the event (e.g., a visit to the doctor's office). This could have resulted in very high levels of recall, especially if the procedures were well known to the parents. By contrast, in many of these same studies, children in the control group received less stressful procedures, but they also were asked to recall highly unfamiliar events for which there was no script and no chance for preevent rehearsal (e.g., meeting a stranger in the school library). Therefore, the recall of the stressed children might have been artificially high and that of the control children artificially low. Even when children received the same procedure, it is possible that few relationships were found because the children were so familiar with these events that there were ceiling effects. In this light, the results of Peters's (1991a, 1991b, 1991c)

studies become particularly interesting because some of his procedures were novel for both stressed and unstressed children, and these events were not known to their parents before participation (e.g., a surprise fire alarm). Thus, the existing literature may conflate the effects of stress with the effects of highly scripted versus unfamiliar events. In terms of age differences in suggestibility, these results indicate that perhaps younger children may be more suggestible than older children in the stressful situations for which they have little understanding of or experience in that situation. At this point in time, these are merely hypotheses in search of data.

Conclusions

In her analyses of appellate court decisions involving expert witnesses' statements about the accuracy of children's testimony, Mason (1991) argued that "courts are unconcerned about scientific knowledge, but are willing to accept the testimony of expert witnesses as long as they have had significant clinical experience" (p. 20). The result is that so-called experts often make claims that are not in accord with—or are even diametrically opposed to—the research we have just reviewed. Experts in psychology, social work, pediatrics, and psychiatry frequently claim in court that children are incapable of lying or are not suggestible (Mason, 1991; McGough, in press). Experts rarely present a careful summary of the research because doing so would probably force them to attenuate their often-strident claims. This can result in distrust of social scientists on grounds that their testimony is speculative at best (Myers, 1987; Zacharias, 1990).

As a way of avoiding this problem, legal jurisdictions traditionally adopted a "Frye test," stipulating that expert testimony is permitted when the research underpinning the expert's testimony is "sufficiently established to have gained general acceptance in the particular field in which it belongs" (*Frye v. the United States*, 1923). Federal Rule 703 has diluted the Frye test standard somewhat, allowing that the database expert witnesses used must be reasonably relied on by experts in the field in forming opinions or inferences on the subject. Hence, it is permissible for experts to disagree about the meaning of research as long as they are aware of it. Because we assume that this review will serve as a basis for the opinions of some who venture into court as expert witnesses, the following three conclusions would seem to meet a traditional Frye test standard.

First and foremost, contrary to the claims made by some (e.g., Melton, 1992), there do appear to be significant age differences in suggestibility, with preschool-aged children being disproportionately more vulnerable to suggestion than either school-aged children or adults. This conclusion follows from a synopsis of the reviewed literature. In approximately 83% (15 out of 18) of the developmental studies that have compared preschoolers with older children or adults, preschoolers were the most suggestible group (see Table 2). To many, this figure of 83% may sound like an overestimate, given the belief that many studies of preschoolers have *not* shown age decrements in suggestibility (e.g., Goodman et al., 1990; Saywitz et al., 1991).⁹ However, reli-

⁹ The figure, 83%, is most likely an underestimate of the extent of actual age differences in suggestibility for three reasons. First, it does

Table 2
Summary of Studies That Compared Suggestibility of Preschoolers to Older Children or to Adults

Study	Reliable age effects for suggestibility
Ceci et al. (in press)	+
Ceci et al. (1987, Exp 1)	+
Ceci et al. (1987, Exp 2)	+
Gordon et al. (1991)	+
Goodman and Aman (1990)	+
Goodman et al. (1990, Exp 4)	+
Goodman, Hirschman, et al. (1991, Exp 3)	+
Goodman, Hirschman, et al. (1991, Exp 2)	+
Goodman, Hirschman, et al. (1991, Exp 4)	+
Goodman and Reed (1986)	+
Oates and Shrimpton (1991)	+
Ornstein et al. (1992)	+
Rudy and Goodman (1991)	+
Saywitz et al. (1991)	+
Marin et al. (1979)	0
Howe (1991)	0
Delamothe and Taplin (1992)	0
Cassel and Bjorklund (1992)	+

Note. To save space, only the first author's or the first and second author's last names appear in the table citations. Exp = experiment. A plus sign denotes that preschoolers were significantly more impaired by misleading questions than older subjects. A zero denotes no significant age differences in the suggestibility effect.

able age differences in suggestibility *were found* in nearly all of these studies despite impressions to the contrary. The reason for this misunderstanding is that those authors sometimes did not focus on age differences in overall suggestibility but instead emphasized a subset of items for which there were no developmental differences in suggestibility, or they focused on a theo-

not reflect the inclusion of studies that report age differences but are not viewed as traditional suggestibility studies. Poole and White's (1991) finding that preschoolers were disproportionately likely to change their answers in response to repeated yes/no questions falls into this class of excluded studies. Second, the 83% estimate reflects a measurement bias, namely, that some studies may have failed to find age differences because of ceiling effects or low statistical power. Third, the denominator for calculating the 83% estimate includes studies that did not use a traditional contrast between a group of children who received misleading information and a control group but instead reported only the far more conservative contrast between the control group performance and a group that received the modified testing procedure (e.g., Delamothe & Taplin, 1992). From all that we know (e.g., Ceci, Ross, & Toglia, 1987), had the traditional contrast been included, such studies probably would have found reliable age differences. Having noted these three factors that may have led to an underestimate of age differences, it is only fair to point out again that we excluded one study (Leippe, Romanczyk, & Manion, 1991) that reported reverse age differences in acquiescence to specific questions (e.g., "Were there flowers on the wall?"), on the ground that such questions traditionally have not been viewed as suggestive, even though they include information not spontaneously provided by the child (Goodman, Rudy, Bottoms, & Aman, 1990, p. 260).

retical issue that was independent of the suggestibility debate. For example, although Goodman and her colleagues chose to focus on segments of their data that did not contain age differences (e.g., abuse-related suggestions, stress induction), it should be kept in mind that they almost always found age differences in overall suggestibility, with the youngest preschoolers being disproportionately more suggestible than older children.

Although the literature clearly reveals age differences in overall suggestibility, the exact mechanisms involved in producing distortion in young children's reports are still being debated by researchers. Until there is a consensus, nothing like a Frye test standard can be met to account for the mechanism by which age differences in suggestibility arise. Some believe that young children's faulty reports are partly the result of an "erasure" mechanism by which erroneous postevent suggestions overwrite or replace the original memory trace (Ceci et al., 1987, 1988; Delamothe & Taplin, 1992; Goodman, Sharma, et al., 1991). By contrast, others believe that social processes lead to distortion in children's reports (Zaragoza et al., 1992). However, all of these researchers agree that regardless of the underlying mechanism, preschoolers' reports *are* distorted by erroneous suggestions.

Controversy exists as to the boundary conditions for younger children's greater suggestibility. Some argue that suggestibility is diminished or even nonexistent when the act in question concerns a central action, when the child is a participant, or when the report is a free narrative (e.g., Fivush, in press; Goodman et al., 1990). The strongest claim of this position is that children are not suggestible concerning central actions, personally experienced, and especially those that involve their own bodies:

Experiments in which children are given misleading information about personally experienced events, as opposed to misinformation about stories, tend to find less of an effect of misleading information. . . . Events which are extremely personally important are probably less prone to suggestion than are less important events. . . . Finally, misleading information is more likely to influence future recall when it is about peripheral details of an event rather than more central aspects. (Fivush, in press, p. 25)

Others, however, have failed to provide support for these claims. Our review of the literature indicates that children can indeed be led to make false or inaccurate reports about very crucial, personally experienced, central events. In Oates and Shrimpton's (1991) study, 4- to 6-year-olds were disproportionately impaired by misleading questions having to do with actions (e.g., having their arms held behind them) compared with older children. By contrast, these investigators failed to find age differences on suggestive questions having to do with nonactions (e.g., a stranger's identity). Similarly, other studies have shown that personally experienced actions are not immune to suggestion. For example, Ceci, Leichtman, Putnick, and Nightingale (1993) have shown that children can be led to falsely report whether they had been kissed while being bathed, and Lindberg (1991) reported that children were suggestible about actions involved in cheating. Recently, Ornstein and his colleagues (Gordon, Ornstein, Clubb, & Nida, 1991) replicated their previous study (Ornstein et al., 1992) on children's memories of their visits to their pediatrician and found that 3-year-old

children were prone to making false claims about "silly events" that involved bodily contact (e.g., Did the nurse lick your knee? Did she blow in your ear?) and that these false claims persisted in repeated interviewing over a 3-month period. Finally, both Cassel and Bjorklund (1992) and Warren and Hagood (in press) found younger children reliably more suggestible about salient, central actions, with fully 42% altering their answer in response to repeated questioning in the former study.

Ironically, studies by Goodman and her colleagues provide some of the most compelling evidence that young children do in fact make false claims about actions, central events, and, even events that could be construed as being sexually abusive (see Goodman et al., 1990). For example, in their anatomical doll study, Goodman and Aman (1990) found that 3- and 5-year-old children frequently gave false answers to abuse-related questions such as "Did he touch your private parts?" (i.e., 32% of 3-year-olds and 24% of 5-year-olds gave inaccurate answers to these types of questions) and to misleading abuse-related questions such as "How many times did he spank you?" (i.e., 24% of 3-year-olds and 3% of 5-year-olds gave inaccurate answers to these types of questions). In addition, these effects are not limited to this one study. In the Rudy and Goodman (1991) study of children's memories of playing with a stranger in a trailer, children made false reports to both misleading abuse questions (i.e., 12% for 4-year-olds and 6% for 7-year-olds) as well as to direct abuse questions (i.e., 18% for 4-year-olds and 10% for 7-year-olds). The age effects were significant only for the direct questions. In the Saywitz et al. (1991) study of children's memories of physical examinations, when all abuse-related questions (nonmisleading and misleading) were pooled, there were significant age differences between 5- and 7-year-old children's reports, and these figures again showed that children do misreport central events, particularly those that involve their own bodies (i.e., error rates were 13% for 5-year-olds and 7% for 7-year-olds). Finally, when 3- to 4-year-olds were interviewed by a neutral interviewer about events surrounding an inoculation, there was an error rate of 23% on questions such as "How many times did she kiss you?" and "She touched your bottom didn't she?" (Goodman et al., 1990, p. 278). That is, many of these children answered yes even though the events did not occur.

The second major conclusion is that contrary to the claims of some, children sometimes lie when the motivational structure is tilted toward lying (e.g., Peters, 1990). In this sense they are probably no different from adults. Because most of the existing data are not developmental, no age comparisons in willingness to lie in response to various motives can be made with confidence at this time, although there is some indication that 5-year-olds are more likely than 3-year-olds to keep secrets (e.g., Bussey, 1992; Pipe & Goodman, in press).

Extreme statements that some have preferred in the media (e.g., children never lie vs. children are incapable of getting it right because they cannot distinguish between reality and fantasy) are not supported by the findings reviewed here. That children are found to lie at times should not surprise anyone except the extreme advocates who have made such baseless claims. More research is needed on age-related shifts in motivational salience to better understand whether different motives, such as fear of reprisals, honoring a promise, or resisting a

bribe, are more or less influential for a given age group. Until such research becomes available, it is safe to conclude that sometimes children will lie, but certainly not all of the time or uniformly in response to all motivational forces.

Third, notwithstanding the aforementioned two points, it is clear that children—even preschoolers—are capable of recalling much that is forensically relevant. That their reports are more vulnerable to distortion than those of older individuals, and that they can be induced to lie in response to certain motives, is not meant to imply that they are incapable of providing accurate testimony. In fact, in most of the studies that have been reported during the past decade, young children were able to accurately recollect the majority of the information they observed, even though they did not recall as much as older children. They may be more likely to succumb to erroneous suggestions than older children, but their vulnerability is a matter of degree only. As was pointed out in the overview, even adults are suggestible (e.g., Belli, 1989; Gudjonsson, 1986; Lindsay, 1990; Loftus, 1979).

Therefore, the question ought not to be whether children are suggestible but whether their level of suggestibility is so much greater than that of an adult as to (a) render them an obstacle to the court's truth-seeking process when they serve as witnesses, (b) require competency hearings to determine whether they ought to be allowed to provide testimony to juries, or (c) require judges to give juries cautionary instructions about their special reliability risks. On the basis of the evidence reviewed in this article, the answer to the first two of these questions appears to be a qualified "no" because, as our review shows, although there are many conditions under which children can report much that is accurate, there are a variety of conditions under which young children appear to be reliably as well as substantively more suggestible than older subjects. Therefore, it is of the utmost importance to examine the conditions prevalent at the time of a child's original report about a criminal event in order to judge the suitability of using that child as a witness in the court. It seems particularly important to know the circumstances under which the initial report of concern was made, how many times the child was questioned, the hypotheses of the interviewers who questioned the child, the kinds of questions the child was asked, and the consistency of the child's report over a period of time. If the child's disclosure was made in a nonthreatening, nonsuggestible atmosphere, if the disclosure was not made after repeated interviews, if the adults who had access to the child prior to his or her testimony are not motivated to distort the child's recollections through relentless and potent suggestions and outright coaching, and if the child's original report remains highly consistent over a period of time, then the young child would be judged to be capable of providing much that is forensically relevant. The absence of any of these conditions would not in and of itself invalidate a child's testimony, but it ought to raise cautions in the mind of the court.

The answer to the third question will depend on whether the courts believe that cautionary instructions about children's reliability risks will serve the useful purpose of taming jurors' unbridled enthusiasm for the young child's credibility (Ross, Dunning, Toglia, & Ceci, 1990; Spencer & Flin, 1990) or whether the

courts believe that such cautionary instructions may exaggerate jurors' preexisting skepticism of children's competencies to a point that is undesirable (Vasek, 1986; also see McCloskey & Eggeth, 1983, for a similar point regarding adult witnesses).

In view of our conclusions, extreme negative opinions about the young child's ability to resist leading questions that have been proffered throughout this century are unwarranted. Assertions from the earlier historical periods, such as "Create, if you will, an idea of what the child is to hear or see, and the child is very likely to see or hear what you desire" (M. R. Brown, 1926, p. 133), are needlessly ungenerous views of children's abilities. Similarly, modern surveys indicating that 91% of psychologists believe that an 8-year-old witness will respond in a way the questioner wished or else say they did not know (Yarmey & Jones, 1983), or the typical college student who believes that children under 6 years of age are "highly prone to be liars, second only to politicians" (Vasek, 1986, p. 157), are also inaccurate.

In light of the full corpus of data that we have reviewed, these extreme opinions are not supported by the available research. This research shows that children are able to encode and retrieve large amounts of information, especially when it is personally experienced and highly meaningful. Equally true, however, is that no good will be served by ignoring that part of the research that demonstrates potentially serious social and cognitive hazards to young child witnesses if adults who have access to them attempt to usurp their memories. Inattention to the full corpus of empirical data will only forestall efforts to improve the way child witnesses are treated and delay needed research into ways of optimizing young children's testimonial accuracy through better interviewing techniques and judicial reform.

References

- Ackerman, B. (1983). Speaker bias in children's evaluation of the external consistency of statements. *Journal of Experimental Child Psychology*, 35, 111-127.
- Andrews, J. A. (1964). The evidence of children. *Criminal Law Review*, 64, 769-777.
- August, R. L., & Forman, B. D. (1989). A comparison of sexually abused and nonsexually abused children's behavioral responses to anatomically correct dolls. *Child Psychiatry and Human Development*, 20, 39-47.
- Aveling, F., & Hargreaves, H. (1921). Suggestibility with and without prestige in children. *British Journal of Psychology*, 11, 53-75.
- Baker-Ward, L., Hess, T., & Flannagan, D. (1990). The effects of involvement on children's memory for events. *Cognitive Development*, 5, 55-69.
- Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice-Hall.
- Baxter, J. (1990). The suggestibility of child witnesses: A review. *Journal of Applied Cognitive Psychology*, 3, 1-15.
- Belli, R. F. (1989). Influences of misleading postevent information: Misinformation interference and acceptance. *Journal of Experimental Psychology: General*, 118, 72-85.
- Belli, R. F., Windschitl, P., McCarthy, T., & Winfrey, S. (1992). Detecting memory impairment with a modified test procedure: Manipulating retention interval with centrally presented event items. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 18, 356-367.
- Benedek, E. P., & Schetky, D. H. (1987). Problems in validating allegations of sexual abuse: Parts 1 and 2. Factors affecting perception and recall of events. *Journal of the American Academy of Child and Adolescent Psychiatry*, 26, 915-922.
- Berliner, L. (1985). The child and the criminal justice system. In A. W. Burgess (Ed.), *Rape and sexual assault* (pp. 199-208). New York: Garland Publishing.
- Binet, A. (1900). *La suggestibilité*. Paris: Schleicher Freres.
- Boat, B., & Everson, M. (1988). The use of anatomical dolls among professionals in sexual abuse evaluations. *Child Abuse and Neglect*, 12, 171-186.
- Boat, B., Everson, M., & Holland, J. (1990). Maternal perceptions of nonabused young children's behaviors after the children's exposure to anatomical dolls. *Child Welfare*, 64, 389-399.
- Bohannon, J. N. (1988). Flashbulb memories for the space shuttle disaster: A tale of two theories. *Cognition*, 29, 179-196.
- Bothwell, R., Deffenbacher, K., & Brigham, J. (1987). Correlation of eyewitness accuracy and confidence: Optimality hypothesis revised. *Journal of Applied Psychology*, 72, 691-695.
- Bottoms, B., Goodman, G., Schwartz-Kenney, B., Sachsenmaier, T., & Thomas, S. (1990, March). *Keeping secrets: Implications for children's testimony*. Paper presented at the biennial meeting of the American Psychology/Law Society, Williamsburg, VA.
- Brainerd, C. J., Kingma, J., & Howe, M. L. (1985). On the development of forgetting. *Child Development*, 56, 1103-1119.
- Brainerd, C. J., & Reyna, V. (1988). Memory loci of suggestibility development: Comment on Ceci, Ross, and Toglia. *Journal of Experimental Psychology: General*, 118, 197-200.
- Brainerd, C. J., Reyna, V. F., Howe, M. L., & Kingma, J. (1990). The development of forgetting and reminiscence. *Monographs of the Society for Research in Child Development*, 55, (3-4, Serial No. 222).
- Brennan, M., & Brennan, R. (1988). *Strange language*. Wagga Wagga, New South Wales, Australia: Riverina Murry Institute of Higher Education.
- Brown, M. R. (1926). *Legal psychology*. New York: Bobbs-Merrill.
- Brown, R., & Kulik, J. (1977). Flashbulb memories. *Cognition*, 5, 73-99.
- Bugental, D., Blue, J., Cortez, V., Fleck, K., & Rodriguez, A. (1992). Influences of witnessed affect on information processing in children. *Child Development*, 63, 774-786.
- Burt, H. E. (1948). *Applied psychology*. Englewood Cliffs, NJ: Prentice Hall.
- Burt, H., & Gaskill, H. (1932). Suggestibility and the form of the question. *Journal of Applied Psychology*, 16, 358-373.
- Bussey, K. (1992). Children's lying and truthfulness: Implications for children's testimony. In S. J. Ceci, M. Leichtman, & M. Putnick (Eds.), *Cognitive and social factors in preschoolers' deception* (pp. 89-110). Hillsdale, NJ: Erlbaum.
- Cassel, W. S., & Bjorklund, D. F. (1992, April). *Age differences and suggestibility of eyewitnesses*. In D. F. Bjorklund & P. A. Ornstein (Coauthors), *Children's memory for real world events: Implications for testimony*. Symposium conducted at the annual meeting of the Conference on Human Development, Atlanta.
- Ceci, S. J. (1990). *On intelligence . . . more or less: A bio-ecological treatise on intellectual development*. Englewood Cliffs, NJ: Prentice-Hall.
- Ceci, S. J. (1991). Some overarching issues in the child suggestibility debate. In J. L. Doris (Ed.), *The suggestibility of children's recollections* (pp. 1-9). Washington, DC: American Psychological Association.
- Ceci, S. J., Caves, R., & Howe, M. J. A. (1981). Children's long term memory for information incongruent with their knowledge. *British Journal of Psychology*, 72, 443-450.

- Ceci, S. J., DeSimone, M., Putnick, M., Lee, J. M., & Toglia, M. (1990, March). *Motives to lie*. Paper presented at the biennial meeting of the American Psychology/Law Society, Williamsburg, VA.
- Ceci, S. J., Leichtman, M., & Putnick, M. (Eds.). (1992). *Cognitive and social factors in early deception*. Hillsdale, NJ: Erlbaum.
- Ceci, S. J., Leichtman, M., Putnick, M., & Nightingale, N. (1993). Age differences in suggestibility. In D. Cicchetti & S. Toth (Eds.), *Child abuse, child development, and social policy*. (pp. 117–137). Norwood, NJ: Ablex.
- Ceci, S. J., Leichtman, M., & White, T. (in press). Interviewing preschoolers: Remembrance of things planted. In D. P. Peters (Ed.), *The child witness: Cognitive, social, and legal issues*. Netherlands: Kluwer.
- Ceci, S. J., Ross, D., & Toglia, M. (1987). Age differences in suggestibility: Psycholegal implications. *Journal of Experimental Psychology: General*, *117*, 38–49.
- Ceci, S. J., Toglia, M., & Ross, D. (1988). On remembering . . . more or less. *Journal of Experimental Psychology: General*, *118*, 250–262.
- Ceci, S. J., Toglia, M., & Ross, D. (1990). The suggestibility of preschoolers' recollections: Historical perspectives on current problems. In R. Fivush & J. Hudson (Eds.), *Knowing and remembering in young children*. (pp. 285–300). New York: Cambridge University Press.
- Chadbourn, J. (1978). *Wigmore on evidence*. Boston: Little, Brown.
- Chandler, M. (1989). Doubt and developing theories of mind. In J. W. Astington, P. Harris, & D. R. Olson (Eds.), *Developing theories of mind* (pp. 387–413). New York: Cambridge University Press.
- Chandler, M., Fritz, A. S., & Hala, S. (1989). Small scale deceit: Deception as a marker of 2-, 3-, and 4-year-olds' early theories of mind. *Child Development*, *60*, 1263–1277.
- Chi, M. T. H., & Ceci, S. J. (1987). Content knowledge and the reorganization of memory. *Advances in Child Development and Behavior*, *20*, 1–37.
- Christianson, S.-Å., & Loftus, E. F. (1987). Memory for traumatic events. *Applied Cognitive Psychology*, *1*, 225–239.
- Clark, H. H., & Clark, E. V. (1977). *Psychology and language*. New York: Harcourt Brace & Jovanovich.
- Clarke-Stewart, A., Thompson, W., & Lepore, S. (1989, May). *Manipulating children's interpretations through interrogation*. Paper presented at the biennial meeting of the Society for Research on Child Development, Kansas City, MO.
- Cohen, R. L. (1975). Children's testimony and hearsay evidence. *Law Reform Commission of Canada: Report on evidence*. Ottawa, Ontario, Canada: Information.
- Cohen, R. L., & Harnick, M. A. (1980). The susceptibility of child witnesses to suggestion. *Law and Human Behavior*, *4*, 201–210.
- Cohn, D. S. (1991). Anatomical doll play of preschoolers referred for sexual abuse and those not referred. *Child Abuse and Neglect*, *15*, 455–466.
- Cole, C. B., & Loftus, E. F. (1987). The memory of children. In S. J. Ceci, M. Toglia, & D. Ross, (Eds.), *Children's eyewitness memory* (pp. 178–208). New York: Springer-Verlag.
- Conte, J. R., Sorenson, E., Fogarty, L., & Rosa, J. D. (1991). Evaluating children's reports of sexual abuse: Results from a survey of professionals. *American Journal of Orthopsychiatry*, *78*, 428–437.
- Cutts, K., & Ceci, S. J. (1988, August). *Memory for cheerios or cheerio memory?* Paper presented at the 96th Annual Convention of the American Psychological Association, Atlanta, GA.
- Dale, P. S., Loftus, E. F., & Rathbun, L. (1978). The influence of the form of the question on the eyewitness testimony of preschool children. *Journal of Psycholinguistic Research*, *7*, 269–277.
- Daro, D., & Mitchel, L. (1990). *Current trends in child abuse reporting and fatalities: The results of the 1989 Annual 50 States Survey*. Washington, DC: National Commission for the Prevention of Child Abuse.
- Davidson, D. (1991, April). *Children's recognitions and recall memory for typical and atypical actions in script-based stories*. Paper presented at the biennial meeting of the Society for Research in Child Development, Seattle, WA.
- Davies, G. M., Tarrant, A., & Flin, R. (1989). Close encounters of a witness kind: Children's memory for a simulated health inspection. *British Journal of Psychology*, *80*, 415–429.
- DeAngelis, T. (1990, September). Ruling tracks brief by APA. *APA Monitor*, pp. 1, 16.
- Deffenbacher, K. A. (1991). A maturing of research on the behaviour of eyewitnesses. *Applied Cognitive Psychology*, *5*, 377–402.
- Delamothe, K., & Taplin, E. (1992, November). *The effect of suggestibility on children's recognition memory*. Paper presented at the Annual Meeting of the Psychonomics Society, St. Louis, MO.
- Dent, H., & Stephenson, G. M. (1979). An experimental study of the effectiveness of different techniques of questioning child witnesses. *British Journal of Social and Clinical Psychology*, *18*, 41–51.
- Duncan, E. M., Whitney, P., & Kunen, S. (1982). Integration of visual and verbal information in children's memories. *Child Development*, *53*, 1215–1223.
- Ervin-Tripp, S. (1978). "Wait for me, Roller Skate!" In S. Ervin-Tripp & C. Mitchell-Kernan (Eds.), *Child discourse* (pp. 165–188). San Diego, CA: Academic Press.
- Everson, M., & Boat, B. (1990). Sexualized doll play among young children: Implications for the use of anatomical dolls in sexual abuse evaluations. *Journal of the American Academy of Child and Adolescent Psychiatry*, *29*, 736–742.
- Farrar, J., & Goodman, G. S. (1990). Developmental differences in the relation between scripts and episodic memory: Do they exist? In R. Fivush & J. Hudson (Eds.), *Knowing and remembering in young children* (pp. 30–64). New York: Cambridge University Press.
- Feher, T. (1988). The alleged molestation victim, the rules of evidence, and the Constitution: Should children really be seen and not heard? *American Journal of Criminal Law*, *14*, 227.
- Fivush, R. (in press). Developmental perspectives on autobiographical recall. In G. S. Goodman & B. Bottoms (Eds.), *Understanding and improving children's testimony*. New York: Guilford Press.
- Fivush, R., & Hamond, N. (1990). Autobiographical memory across the preschool years: Toward reconceptualizing childhood amnesia. In R. Fivush & J. Hudson (Eds.), *Knowing and remembering in young children* (pp. 223–248). New York: Cambridge University Press.
- Flanagan, O. (1992). Other minds, obligations, and honesty. In S. J. Ceci, M. Leichtman, & M. Putnick (Eds.), *Cognitive and social factors in early deception*. (pp. 111–126). Hillsdale, NJ: Erlbaum.
- Flavell, J. H. (1985). *Cognitive development* (2nd ed.). Englewood Cliffs, NJ: Prentice-Hall.
- Flavell, J. H., Flavell, E., & Green, F. L. (1987). Young children's knowledge about the apparent-real and pretend-real distinctions. *Developmental Psychology*, *23*, 816–822.
- Flin, R., Boon, J., Knox, A., & Bull, R. (1992). Children's memories following a five-month delay. *British Journal of Psychology*, *83*, 323–336.
- Foley, M. A., & Johnson, M. K. (1985). Confusions between memories for performed and imagined actions. *Child Development*, *56*, 1145–1155.
- Foley, M. A., Johnson, M. K., & Raye, C. L. (1983). Age-related confusion between memories for thoughts and memories for speech. *Child Development*, *54*, 51–60.
- Foley, M. A., Santini, C., & Sopaşakis, M. (in press). Discriminating between memories: Evidence for children's spontaneous elaborations. *Journal of Experimental Child Psychology*.
- Freiberg, P. (1990, December). Jurors perceive when children lie about abuse. *APA Monitor*, p. 32.

- Freud, S. (1966). New introductory lectures on psychoanalysis. In J. Strachey (Ed. and Trans.), *The standard edition of the complete psychological works of Sigmund Freud* (Vol. 15-16). London: Hogarth Press. (Original work published 1933)
- Frye v. the United States, 293 F.1013 (DC. Cir. 1923).
- Garbarino, J., & Stott, F. M. (1989). *What children can tell us*. San Francisco: Jossey-Bass.
- Gardner, R. (1989). *Sex abuse hysteria: Salem Witch Trials revisited*. Longwood, NJ: Creative Therapeutics Press.
- Garvey, C. (1984). *Children's talk*. Cambridge, MA: Harvard University Press.
- Geiselman, R., Saywitz, K., & Bornstein, G. (1990). *Effects of cognitive interviewing, practice, and interview style on children's recall performance*. Unpublished manuscript.
- Gelman, R., Meck, E., & Merkin, S. (1986). Young children's numerical competence. *Cognitive Development, 1*, 1-29.
- Glaser, D., & Collins, C. (1989). The response of young, non-sexually abused children to anatomically correct dolls. *Journal of Child Psychology and Psychiatry, 30*, 547-560.
- Gold, P. (1987). Sweet memories. *American Scientist, 75*, 151-155.
- Goodman, G. (1984a). Children's testimony in historical perspective. *Journal of Social Issues, 40*, 9-31.
- Goodman, G. (1984b). The child witness: Conclusions and future directions. *Journal of Social Issues, 40*, 157-175.
- Goodman, G. S. (1990, August). *Psychology and the media*. In D. Singer (Chair), Symposium conducted at the 98th Annual Convention of the American Psychological Association, Boston.
- Goodman, G. S. (1991). Commentary: On stress and accuracy in research on children's testimony. In J. Doris (Ed.), *The suggestibility of children's recollections* (pp. 77-82). Washington, DC: American Psychological Association.
- Goodman, G., & Aman, C. (1990). Children's use of anatomically detailed dolls to recount an event. *Child Development, 61*, 1859-1871.
- Goodman, G. S., & Clarke-Stewart, A. (1991). Suggestibility in children's testimony: Implications for child sexual abuse investigations. In J. L. Doris (Ed.), *The suggestibility of children's recollections* (pp. 92-105). Washington, DC: American Psychological Association.
- Goodman, G. S., Hirschman, J. E., Hepps, D., & Rudy, L. (1991). Children's memory for stressful events. *Merrill Palmer Quarterly, 37*, 109-158.
- Goodman, G. S., Levine, M., Melton, G., & Ogden, D. (1991). Child witnesses and the confrontation clause: The American Psychological Association Brief in *Maryland v. Craig*. *Law and Human Behavior, 15*, 13-30.
- Goodman, G. S., & Reed, R. S. (1986). Age differences in eyewitness testimony. *Law and Human Behavior, 10*, 317-332.
- Goodman, G. S., Rudy, L., Bottoms, B., & Aman, C. (1990). Children's concerns and memory: Issues of ecological validity in the study of children's eyewitness testimony. In R. Fivush & J. Hudson (Eds.), *Knowing and remembering in young children* (pp. 249-284). New York: Cambridge University Press.
- Goodman, G. S., Sharma, A., Golden, M., & Thomas, S. (1991, April). *The effects of mothers' and strangers' interviewing strategies on children's reporting of real-life events*. Paper presented at the biennial meeting of the Society for Research in Child Development, Seattle, WA.
- Goodman, G. S., Wilson, M. E., Hazan, C., & Reed, R. S. (1989, April). *Children's testimony nearly four years after an event*. Paper presented at the annual meeting of the Eastern Psychological Association, Boston.
- Gordon, B. N., Ornstein, P. A., Clubb, P. A., & Nida, R. E. (1991, November). *Visiting the pediatrician: Long-term retention and forgetting*. Paper presented at the annual meeting of Psychonomic Society, San Francisco.
- Grice, H. P. (1975). Logic and conversation. In P. Cole & J. L. Morgan (Eds.), *Syntax and semantics* (pp. 41-58). San Diego, CA: Academic Press.
- Gudjonsson, G. (1986). The relationship between interrogative suggestibility and acquiescence: Empirical findings and theoretical implications. *Personality and Individual Differences, 7*, 195-199.
- Harris, P., Brown, E., Marriott, C., Whittall, S., & Harmer, S. (1991). Monsters, ghosts and witches: Testing the limits of the fantasy-reality distinction in young children. *British Journal of Developmental Psychology, 9*, 105-123.
- Haugaard, J., & Repucci, N. D. (1992). Children and the truth. In S. J. Ceci, M. Leichtman, & M. Putnick (Eds.), *Social and cognitive factors in early deception*. (pp. 29-46). Hillsdale, NJ: Erlbaum.
- Hollingsworth, J. (1986). *Unspeakable acts*. Chicago: Congdon & Weed.
- Howe, M. L. (1991). Misleading children's story recall: Forgetting and reminiscence of the facts. *Developmental Psychology, 27*, 746-762.
- Hudson, J. (1990). Constructive processes in children's event memories. *Developmental Psychology, 26*, 180-187.
- Hudson, J., & Nelson, K. (1986). Repeated encounters of a similar kind: Effects of familiarity on children's autobiographic memory. *Cognitive Development, 1*, 253-271.
- Hughes, M., & Grieve, R. (1980). On asking children bizarre questions. *First Language, 1*, 149-160.
- Hurllock, E. (1930). Suggestibility in children. *Journal of Genetic Psychology, 37*, 59-74.
- Jampole, L., & Weber, M. K. (1987). An assessment of the behavior of sexually abused and nonsexually abused children with anatomically correct dolls. *Child Abuse and Neglect, 11*, 187-192.
- Johnson, M. K. (1991). Reality monitoring: Evidence from confabulation in organic brain disease patients. In G. Prigatano & D. Schacter (Eds.), *Awareness of deficit after brain injury* (pp. 124-140). New York: Oxford University Press.
- Johnson, M. K., & Raye, C. L. (1981). Reality monitoring. *Psychological Review, 88*, 67-85.
- Jones, D., & McGraw, J. M. (1987). Reliable and fictitious accounts of sexual abuse in children. *Journal of Interpersonal Violence, 2*, 27-45.
- Jones, D. C., Swift, D. J., & Johnson, M. (1988). Nondeliberate memory for a novelty event among preschoolers. *Developmental Psychology, 24*, 641-645.
- Kail, R. V. (1989). *The development of memory in children* (2nd ed.). New York: Freeman.
- King, M., & Yuille, J. (1987). Suggestibility and the child witness. In S. J. Ceci, M. Toglia, & D. Ross (Eds.), *Children's eyewitness memory* (pp. 24-35). New York: Springer-Verlag.
- Kramer, T., Buckhout, R., Fox, P., Widman, E., & Tusche, B. (1991). Effects of stress on recall. *Applied Cognitive Psychology, 5*, 483-488.
- Leekam, S. (1992). Believing and deceiving: Steps to becoming a good liar. In S. J. Ceci, M. Leichtman, & M. Putnick (Eds.), *Social and cognitive factors in early deception*. (pp. 47-62). Hillsdale, NJ: Erlbaum.
- Leippe, M. R., Romanczyk, A., & Manion, A. P. (1991). Eyewitness memory for a touching experience: Accuracy in eyewitness identification. *Journal of Applied Psychology, 76*, 367-379.
- Lewis, M., Stranger, C., & Sullivan, M. (1989). Deception in three-year-olds. *Developmental Psychology, 25*, 439-443.
- Lindberg, M. (1980). Is knowledge base development a necessary and sufficient condition for memory development? *Journal of Experimental Child Psychology, 30*, 401-410.
- Lindberg, M. (1991). A taxonomy of suggestibility and eyewitness memory: Age, memory process, and focus of analysis. In J. L. Doris

- (Ed.), *The suggestibility of children's recollections* (pp. 47–55). Washington, DC: American Psychological Association.
- Lindsay, D. S. (1990). Misleading suggestions can impair eyewitnesses' ability to remember event details. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*, 1077–1083.
- Lindsay, D. S., Johnson, M. K., & Kwon, P. (1991). Developmental changes in memory source monitoring. *Journal of Experimental Child Psychology*, *52*, 297–318.
- Linton, M. (1982). Transformation of memory in everyday life. In U. Neisser (Ed.), *Memory observed: Remembering in natural contexts* (pp. 77–91). New York: W. H. Freeman.
- Lipmann, O. (1911). Pedagogical psychology of report. *Journal of Educational Psychology*, *2*, 253–261.
- Lipscomb, T. J., Bregman, N., & McCallister, H. A. (1984). A developmental inquiry into the effects of postevent information on eyewitness accounts. *Journal of Genetic Psychology*, *146*, 551–556.
- List, J. A. (1986). Age and schematic differences in the reliability of eyewitness testimony. *Developmental Psychology*, *22*, 50–57.
- Loftus, E. F. (1979). *Eyewitness testimony*. Cambridge, MA: Harvard University Press.
- Loftus, E. F. (1986). Ten years in the life of an expert witness. *Law and Human Behavior*, *10*, 241–263.
- Loftus, E. F., & Ceci, S. J. (1991). Research findings: What do they mean? In J. L. Doris (Ed.), *The suggestibility of children's recollections* (pp. 129–133). Washington, DC: American Psychological Association.
- Loftus, E. F., Miller, D., & Burns, H. (1978). Semantic integration of verbal information into a visual memory. *Journal of Experimental Psychology: Human Learning and Memory*, *4*, 19–31.
- Loftus, E. F., & Palmer, J. C. (1974). Reconstruction of automobile destruction: An example of the interaction between language and memory. *Journal of Verbal Learning and Verbal Behavior*, *13*, 585–589.
- Loftus, E. F., & Zanni, G. (1975). Eyewitness testimony: The influence of wording of a question. *Bulletin of the Psychonomic Society*, *5*, 86–88.
- Loh, W. D. (1981). Psychological research: Past and present. *Michigan Law Review*, *79*, 659–707.
- Manshel, L. (1990). *Nap time*. New York: Kensington Publishing.
- Marin, B. V., Holmes, D. L., Guth, M., & Kovac, P. (1979). The potential of children as eyewitnesses. *Law and Human Behavior*, *3*, 295–305.
- Martin, C. L., & Halverson, C. F. (1983). The effects of sex-typing schemas on young children's memory. *Child Development*, *54*, 563–574.
- Maryland v. Craig, 110 S. Ct. 3157, 3169 (1990).
- Mason, M. A. (1991, Fall–Winter). A judicial dilemma: Expert witness testimony in child sex abuse trials. *Psychiatry & Law*, 185–219.
- Masson, J. M. (1984, February). Freud and the seduction theory. *Atlantic Monthly*, pp. 33–60.
- McCloskey, M., & Egeth, H. W. (1983). Eyewitness identification: What can a psychologist tell a jury? *American Psychologist*, *38*, 550–563.
- McCloskey, M., Wible, C., & Cohen, N. J. (1988). Is there a special flashbulb-memory mechanism? *Journal of Experimental Psychology: General*, *117*, 171–181.
- McCloskey, M., & Zaragoza, M. (1985a). Misleading postevent information and memory for events: Arguments and evidence against the memory impairment hypothesis. *Journal of Experimental Psychology: General*, *114*, 1–16.
- McCloskey, M., & Zaragoza, M. (1985b). Postevent information and memory: Reply to Loftus, Schooler, and Wagenaar. *Journal of Experimental Psychology: General*, *114*, 381–387.
- McConnell, T. R. (1963). Suggestibility in children as a function of chronological age. *Journal of Abnormal and Social Psychology*, *67*, 286–289.
- McGough, L. (in press). *Fragile voices: The child witness in American courts*. New Haven, CT: Yale University Press.
- Melton, G. (1992). Children as partners for justice: Next steps for developmentalists. *Monographs of the Society for Research in Child Development*, *57*, (5, Serial No. 229).
- Melton, G., & Thomson, R. (1987). Getting out of a rut: Detours to less travelled paths in child witness research. In S. J. Ceci, M. Toglia, & D. Ross (Eds.), *Children's eyewitness memory* (pp. 209–229). New York: Springer-Verlag.
- Messerschmidt, R. (1933). The suggestibility of boys and girls between the ages of six and sixteen. *Journal of Genetic Psychology*, *43*, 422–437.
- Mitchell, R. W., & Thompson, N. S. (Eds.) (1986). *Deception: Perspectives on human and non-human deceit*. Albany: State University of New York Press.
- Moore, C. (1907). Yellow psychology. *Law Notes*, *11*, 125–127.
- Moore, C. C. (1908). Psychology in the courts. *Law Notes*, *11*, 185–187.
- Morison, P., & Gardner, H. (1978). Dragons and dinosaurs: The child's capacity to differentiate fantasy from reality. *Child Development*, *49*, 642–648.
- Moss, D. C. (1988). "Real" dolls too suggestive: Do anatomically correct dolls lead to false abuse charges? *American Bar Association Journal*, *24* (Dec. 1).
- Moston, S. (1987). The suggestibility of children in interview studies. *First Language*, *7*, 67–78.
- Münsternberg, H. (1907a, April). Nothing but the truth. *McClure's Magazine*, pp. 26–29.
- Münsternberg, H. (1907b). Yellow psychology: Dr. Münsterberg replies to Mr. Moore. *Law Notes*, *11*, 145–146.
- Myers, J. B. (1987). *Child witness law and practice*. New York: Wiley.
- Myles-Worsley, M., Cromer, C., & Dodd, D. (1986). Children's preschool script construction: Reliance on general knowledge after memory fades. *Developmental Psychology*, *22*, 22–30.
- Nathan, D. (1993). Revisiting Country Walk. *Issues in Child Abuse Allegations*, *5*, 1–11.
- Nelson, K., & Gruendel, J. (1979). At morning it's lunchtime: A scriptal view of children's dialogues. *Discourse Processes*, *2*, 73–94.
- Nurcombe, B. (1986). The child as witness: Competency and credibility. *Journal of the American Academy of Child Psychiatry*, *25*, 473–480.
- Oates, K., & Shrimpton, S. (1991). Children's memories for stressful and non-stressful events. *Medicine, Science, and the Law*, *31*, 4–10.
- Ornstein, P. A. (Ed.). (1978). *Memory development in children*. Hillsdale, NJ: Erlbaum.
- Ornstein, P. A., Gordon, B. N., & Larus, D. (1992). Children's memory for a personally experienced event: Implications for testimony. *Applied Cognitive Psychology*, *6*, 49–60.
- Otis, M. (1924). A study of suggestibility in children. *Archives of Psychology*, *11*, 5–108.
- Perner, J. (1991). *Understanding the representational mind*. Cambridge, MA: MIT Press.
- Perris, E. E., Myers, N. A., & Clifton, R. K. (1990). Long term memory for a single infancy experience. *Child Development*, *61*, 1796–1807.
- Peters, D. P. (1990, March). *Confrontational stress and lying*. Paper presented at the biennial meeting of the American Psychology/Law Society, Williamsburg, VA.
- Peters, D. P. (1991a). The influence of stress and arousal on the child witness. In J. L. Doris (Ed.), *The suggestibility of children's recollections* (pp. 60–76). Washington, DC: American Psychological Association.
- Peters, D. P. (1991b). Commentary: Response to Goodman. In J. L.

- Doris (Ed.), *The suggestibility of children's recollections* (pp. 86–91). Washington, DC: American Psychological Association.
- Peters, D. P. (1991c, April). *Confrontational stress and children's testimony*. Paper presented at the biennial meeting of the Society for Research in Child Development, Seattle, WA.
- Pettit, F., Fegan, M., & Howie, P. (1990, September). *Interviewer effects on children's testimony*. Paper presented at the International Congress on Child Abuse and Neglect, Hamburg, Federal Republic of Germany.
- Piaget, J. (1926). *The language and thought of the child*. London: Routledge & Kegan Paul.
- Pillemer, D. B. (1990). Clarifying the flashbulb memory concept: Comment on McCloskey, Wible, and Cohen (1988). *Journal of Experimental Psychology: General*, *119*, 92–96.
- Pillemer, D. B. (in press). Preschool children's memories of personal circumstances: The fire alarm study. In E. Winograd & U. Neisser (Eds.), *Affect and accuracy in recall: The problem of flashbulb memories*. New York: Cambridge University Press.
- Pillemer, D. B., Koff, E., Rhinehart, E. D., & Rierdan, J. (1987). Flashbulb memories of menarche and adult menstrual distress. *Journal of Adolescence*, *10*, 187–199.
- Pipe, M. E., & Goodman, G. S. (in press). Elements of secrecy: Implications for children's testimony. *Behavioral Sciences and the Law*.
- Poole, D. A. (in press). Two years later: Effects of question repetition and retention interval on the eyewitness testimony of children and adults. *Developmental Psychology*.
- Poole, D. A., & White, L. T. (1991). Effects of question repetition on the eyewitness testimony of children and adults. *Developmental Psychology*, *27*, 975–986.
- Powers, P., Andriks, J. L., & Loftus, E. F. (1979). Eyewitness accounts of females and males. *Journal of Applied Psychology*, *64*, 339–347.
- Pratt, C. (1991). On asking children—and adults—bizarre questions. *First Language*, *10*, 167–175.
- Raskin, D., & Yuille, J. (1989). Problems in evaluating interviews of children in sexual abuse cases. In S. J. Ceci, M. Toglia, & D. F. Ross (Eds.), *Adults' perceptions of children's testimony* (pp. 184–207). New York: Springer-Verlag.
- Read, B., & Cherry, L. (1978). Preschool children's productions of directive forms. *Discourse Processes*, *1*, 233–245.
- 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*, 743–746.
- Remmers, H., Cutler, M., & Jones, P. (1940). Waking suggestibility in children: General or specific? *Journal of Genetic Psychology*, *56*, 87–93.
- Romaine, S. (1984). *The language of children and adolescents*. Cambridge, MA: Basil Blackwell.
- Rose, S., & Blank, M. (1974). The potency of context in children's cognition: An illustration through conservation. *Child Development*, *45*, 499–502.
- Ross, D. F., Dunning, D., Toglia, M., & Ceci, S. J. (1990). The child in the eyes of the jury. *Law and Human Behavior*, *14*, 5–23.
- Rovee-Collier, C., & Borza, M. A. (in press). Infants' eyewitness testimony: Integrating postevent information into a prior memory representation. *Memory & Cognition*.
- Rudy, L., & Goodman, G. S. (1991). Effects of participation on children's reports: Implications for children's testimony. *Developmental Psychology*, *27*, 527–538.
- Saywitz, K. (1987). Children's testimony: Age-related patterns of memory errors. In S. J. Ceci, M. Toglia, & D. Ross (Eds.), *Children's eyewitness memory* (pp. 36–52). New York: Springer-Verlag.
- Saywitz, K. (1989). Children's conceptions of the legal system: "Court is a place to play basketball." In S. J. Ceci, M. Toglia, & D. Ross (Eds.), *Perspectives on children's testimony* (pp. 131–157). New York: Springer-Verlag.
- Saywitz, K., Geiselman, R., & Bornstein, G. (1992). Effects of cognitive interviewing, practice, and interview style on children's recall performance. *Journal of Applied Psychology*, *77*, 744–756.
- Saywitz, K., Goodman, G., Nicholas, G., & Moan, S. (1991). Children's memory for genital exam: Implications for child sexual abuse. *Journal of Consulting and Clinical Psychology*, *59*, 682–691.
- Schneider, W., & Pressley, M. (1989). *Memory development between 2 and 20*. New York: Springer-Verlag.
- Schuman, D. C. (1986). False allegations of physical and sexual abuse. *Bulletin of the American Academy of Psychiatry and the Law*, *14*, 5–21.
- Schwartz-Kenney, B. M., & Goodman, G. S. (1991, April). *Effects of misinformation on children's memory for a real-life event*. Paper presented at the biennial meeting of the Society for Research in Child Development, Seattle, WA.
- Sherman, I. (1925). The suggestibility of normal and mentally defective children. *Comparative Psychology Monographs*, No. 2.
- Siegal, M., Waters, L., & Dinwiddy, L. (1988). Misleading children: Causal attributions for inconsistency under repeated questioning. *Journal of Experimental Child Psychology*, *45*, 438–456.
- Sivan, A. B., Schor, D. P., Koepl, G. K., & Noble, L. D. (1988). Interaction of normal children with anatomical dolls. *Child Abuse and Neglect*, *12*, 295–304.
- Sodian, B. (1991). The development of deception in young children. *British Journal of Developmental Psychology*, *9*, 173–188.
- Sonnenschein, S., & Whitehurst, G. (1980). The development of communication: When a bad model makes a good teacher. *Journal of Experimental Child Psychology*, *3*, 371–390.
- Spencer, J., & Flin, R. (1990). *The evidence of children: The law and the psychology*. London: Blackstone.
- Stern, W. (1910). Abstracts of lectures on the psychology of testimony and on the study of individuality. *American Journal of Psychology*, *21*, 270–282.
- Steward, M. (1989). *The development of a model interview for young child victims of sexual abuse*. (Tech. Rep. No. 90CA1332). Washington, DC: U.S. Department of Health and Human Services.
- Stouthamer-Loeber, M. (1987, April). *Mothers' perceptions of children's lying and its relationship to behavior problems*. Paper presented at the biennial meeting of the Society for Research in Child Development, Baltimore, MD.
- Tate, C. S., & Warren-Leubecker, A. R. (1990, March). Can young children lie convincingly if coached by adults? In S. J. Ceci (Chair), *Do children lie? Narrowing the uncertainties*. Symposium conducted at the biennial meeting of the American Psychology/Law Society, Williamsburg, VA.
- Tate, C. S., Warren, A. R., & Hess, T. H. (1992). Adults' liability for children's "lie-ability": Can adults coach children to lie successfully? In S. J. Ceci, M. Leichtman, & M. Putnick (Eds.), *Cognitive and social factors in early deception*. (pp. 69–88). Hillsdale, NJ: Erlbaum.
- Taylor, R., & Howell, M. (1973). The ability of 3-, 4-, and 5-year-olds to distinguish fantasy from reality. *Journal of Genetic Psychology*, *122*, 315–318.
- Terr, L. (1983). Chowchilla revisited: The effects of psychic trauma four years after a schoolbus kidnapping. *American Journal of Psychiatry*, *140*, 1543–1550.
- Terr, L. (1988). Anatomically correct dolls: Should they be used as a basis for expert testimony? *Journal of the American Academy of Child and Adolescent Psychiatry*, *27*, 254–257.
- Terr, L. (1990). *Too scared to cry: Psychic trauma in childhood*. New York: Harper & Row.

- Toglia, M. (1991). Commentary: Memory impairment—It is more common than you think. In J. L. Doris (Ed.), *The suggestibility of children's recollections* (pp. 40–46). Washington, DC: American Psychological Association.
- Tucker, A., Merton, P., & Luszcz, M. (1990). The effect of repeated interviews on young children's eyewitness memory. *Australian and New Zealand Journal of Criminology*, 23, 117–123.
- Tulving, E. E., & Watkins, M. J. (1975). Structure of memory traces. *Psychological Review*, 82, 261–275.
- Underwager, R., & Wakefield, H. (1990). *The real world of child interrogations*. Springfield, IL: Charles C Thomas.
- Vandermaas, M. (1991, April). *Assessment of young children's anxiety during dental procedures*. Paper presented at the biennial meeting of the Society for Research in Child Development, Seattle, WA.
- Varendonck, J. (1911). Les témoignages d'enfants dans un proces retentissant [The testimony of children in a famous trial]. *Archives de Psychologie*, 11, 129–171.
- Vasek, M. E. (1986). Lying as a skill: The development of deception in children. In R. W. Mitchell & N. Thompson (Eds.), *Deception: Perspectives on human and nonhuman deceit* (pp. 144–168). New York: State University of New York: Press.
- Walker, A. G. (1988, June). *Questioning young children in court*. Paper presented at the Law and Society annual meeting, Washington, DC.
- Warren, A., & Hagood, P. (in press). Effects of timing and type of questioning on eyewitness accuracy and suggestibility. In M. Zaragoza (Ed.), *Memory, suggestibility, and eyewitness testimony in children and adults*. New York: Hemisphere.
- Warren, A. R., Hulse-Trotter, K., & Tubbs, E. (1991). Inducing resistance to suggestibility in children. *Law and Human Behavior*, 15, 273–285.
- Warren, A. R., & Swartwood, J. N. (in press). Developmental issues in flashbulb memory research: Children recall the *Challenger* event. In E. Winograd & U. Neisser (Eds.), *Affect and accuracy in recall: The problem of flashbulb memories*. New York: Cambridge University Press.
- Warren-Leubecker, A. R., Tate, C., Hinton, I., & Ozbek, N. (1989). What do children know about the legal system and when do they know it? First steps down a less travelled path in child witness research. In S. J. Ceci, M. Toglia, & D. Ross (Eds.), *Perspectives on children's testimony* (pp. 158–183). New York: Springer-Verlag.
- Whipple, G. M. (1909). The observer as reporter: A survey of the "psychology of testimony." *Psychological Bulletin*, 6, 153–170.
- Whipple, G. M. (1911). The psychology of testimony. *Psychological Bulletin*, 8, 307–309.
- Whipple, G. M. (1912). Psychology of testimony and report. *Psychological Bulletin*, 9, 264–269.
- Whipple, G. M. (1913). Psychology of testimony and report. *Psychological Bulletin*, 10, 264–268.
- White, S., Strom, G., Santili, G., & Halpin, B. M. (1986). Interviewing young sexual abuse victims with anatomically correct. *Child Abuse and Neglect*, 10, 519–530.
- Whiten, A. (1991). *Natural theories of mind: The evolution, development, and simulation of everyday mindreading*. Oxford, England: Basil Blackwell.
- Wigmore, J. H. (1909). *Evidence*. Boston: Little, Brown.
- Wigmore, J. H. (1935). *Evidence in trials at common law* (Vol. 6). Boston: Little, Brown.
- Wilson, J. C., & Pipe, M. E. (1989). The effects of cues on young children's recall of real events. *New Zealand Journal of Psychology*, 18, 65–70.
- Winograd, E., & Killinger, W. A., Jr. (1983). Relating age at encoding in early childhood to adult recall: Development of flashbulb memories. *Journal of Experimental Psychology: General*, 112, 413–422.
- Wolfner, G., Faust, D., & Dawes, R. (1993). The use of anatomically detailed dolls in sexual abuse evaluations: The state of the science. *Applied and Preventive Psychology*, 2, 1–11.
- Yarmey, D., & Jones, J. (1983). Is the psychology of eyewitness identification a matter of common sense? In S. Lloyd & B. Clifford (Eds.), *Evaluating witness evidence* (pp. 13–40). New York: Wiley.
- Zacharias, F. (1990). Rethinking confidentiality: II. Is confidentiality constitutional? *Iowa Law Review*, 75, 601.
- Zaragoza, M. (1991). Preschool children's susceptibility to memory impairment. In J. L. Doris (Ed.), *The suggestibility of children's recollections* (pp. 27–39). Washington, DC: American Psychological Association.
- Zaragoza, M., Dahlgren, D., & Muench, J. (1992). The role of memory impairment in children's suggestibility. In M. L. Howe, C. J. Brainerd, & V. F. Reyna (Eds.), *Development of long term retention*. (pp. 184–216). New York: Springer-Verlag.

Received May 30, 1991

Revision received May 20, 1992

Accepted May 20, 1992 ■