

Do Jurors “Know” What Isn’t So About Child Witnesses?

Jodi A. Quas,^{1,3} William C. Thompson,² and K. Alison Clarke-Stewart¹

Are expert witnesses needed in child sexual abuse cases to educate jurors about children’s memory, suggestibility, and reactions to abuse, or do jurors already know what such experts could tell them? To cast light on this question, we surveyed jurors and jury-eligible college students and compared their beliefs with what is known via scientific research regarding children’s memory and ability to testify, reactions to interrogation, and reactions to sexual abuse. We also asked participants to infer results of four widely cited studies of children’s suggestibility. Participants’ beliefs were consistent with findings from research on some issues (e.g., that children can be led to claim that false events occurred) but diverged from the scientific consensus on other issues (e.g., whether children can remember painful events in infancy). Similarly, participants sometimes overestimated and sometimes underestimated the level of suggestibility observed in empirical studies. Individual differences in accuracy were related to participants’ gender, education and ethnicity, and there was considerable disagreement among participants on many questions. Implications of findings for the admissibility of expert testimony in child abuse cases are discussed.

KEY WORDS: children; expert testimony; eyewitness testimony; juror decision making; suggestibility.

INTRODUCTION

Each year, thousands of children are interviewed about suspected abuse and neglect. Children’s statements during these interviews often play a crucial role in subsequent legal proceedings, helping to decide the fate of children as well as the individuals accused of committing the abusive acts. Consequently, the reliability of children’s accounts of maltreatment has been a matter of enormous concern.

Academic psychologists responded to this concern by conducting several hundred empirical studies on the accuracy of children’s memory, children’s sug-

¹Department of Psychology and Social Behavior, University of California, Irvine, California.

²Department of Criminology, Law and Society, University of California, Irvine, California.

³To whom correspondence should be addressed at Department of Psychology and Social Behavior, University of California, 3340 Social Ecology II, Irvine, California 92697-7085; e-mail: jquas@uci.edu.

gestibility, and variables that influence the reliability and completeness of children's statements about their experiences (see Bruck, Ceci, & Hembrooke, 2002; Quas, Goodman, Ghetti, & Redlich, 2000, for reviews). This large body of research has greatly advanced scientific knowledge and has produced consensus among experts on a number of key issues (Ceci & Bruck, 1993; Kovera & Borgida, 1997; Morison & Greene, 1992). However, decisions about whether to believe children are often made by individuals unfamiliar with the scientific literature, most often jurors chosen to be representative of the community. A critical question thus concerns whether jurors have adequate knowledge of children's memory, suggestibility, and eyewitness abilities to make good decisions, or whether it would be helpful to inform jurors, through expert testimony, about results of scientific studies on children's reliability as witnesses.

This question of whether jurors would benefit from expert witness testimony on the topic of children's suggestibility has been debated among academic scholars (e.g., Ceci & Friedman, 2000; Lyon, 1999; McAuliff & Kovera, 2005; Myers, 1997) and in the courts (e.g., *State v. Sloan*, 1995). Some commentators have called for expert testimony to educate jurors about scientific findings, specifically to describe literature concerning the strengths and limitations of children's performance as witnesses, information that goes beyond the layperson's commonsense knowledge. Other commentators have argued that expert testimony is unnecessary because jurors already know enough to evaluate children's testimony. The present study was designed to inform this debate. Specifically, we investigated what actual and potential jurors know—and think they know—about child witnesses. By studying the extent to which jurors' understanding of children's reliability corresponds to, or diverges from, findings in the scientific literature, we hope to cast light on the need for expert testimony in legal cases involving child witnesses.

Social Framework Evidence Regarding Child Witnesses

Expert testimony about children's ability to recount prior experiences and about factors that affect children's memory accuracy is an example of what Monahan and Walker (1988) called "social framework evidence." This testimony is offered to provide the trier-of-fact (usually a jury) with knowledge derived from scientific research that gives an appropriate context for understanding specific evidence in a given case. Social framework testimony is distinguished from expert testimony based on direct examination of the child, for example by a therapist or clinician. An expert need not have examined the child to provide social framework evidence. Instead, the expert may simply report findings and conclusions from the scientific literature that are relevant to the jury's evaluation of the child's statements, without expressing an opinion about the child's credibility.

It appears that prosecutors were the first to present social framework testimony about children's reliability as witnesses (Cacciola, 1986; McCord, 1986; Myers et al., 1989; Morison & Greene, 1992; Roe, 1985; Summit, 1983). This testimony was designed to bolster children's credibility by rebutting certain misconceptions about children that, it was feared, jurors might hold; for example, that children have difficulty distinguishing fact from fantasy, that delays in reporting signal unreliable

claims, and that truthful claims are never recanted. Several legal scholars (Myers, 1993; Myers et al., 1989; Roe, 1985) advocated the use of such expert testimony to inform jurors about psychological studies indicating that children above 3 or 4 years of age are capable of distinguishing fact from fantasy (e.g., Johnson & Foley, 1984), are not unduly suggestible (e.g., Brigham, Van Verst, & Bothwell, 1981; Rudy & Goodman, 1991), often delay reporting abuse (e.g., Kellogg & Huston, 1995; Elliot & Briere, 1994), sometimes recant truthful charges (e.g., Elliot & Briere, 1994; Gordon & Jaudes, 1996), and are capable of providing accurate accounts of salient, personal experiences (e.g., Fivush, Hudson, & Nelson, 1984; Saywitz, Goodman, Nicholas, & Moan, 1991).

This type of expert testimony sparked some criticism (e.g., Cirelli, 1991; Flint, 1995). Some commentators suggested that the testimony failed to meet conventional standards for the admissibility of scientific evidence (Levy, 1989) and that partisan experts overstated the evidence (Underwager & Wakefield, 1992). For example, Levy (1989) complained of prosecution experts who were "fiercely committed to the role of children's advocate, devoted to preferring child protection to any other value" (p. 396).

During the 1990s, a new wave of studies emerged that focused on children's reactions to suggestive interview techniques (e.g., Ceci, Loftus, Leichtman, & Bruck, 1994; Poole & Lindsay, 1995; Thompson, Clarke-Stewart, & Lepore, 1997). These studies clearly demonstrated that children's accounts of events can be radically altered through the use of highly suggestive interrogation practices. In conjunction with these new studies, defense attorneys began presenting expert testimony to help jurors evaluate the credibility of children who had been (or might have been) interviewed in a suggestive manner. Thus, social framework testimony about children's accuracy was no longer exclusively the tool of prosecutors seeking to bolster children's accounts—it began to be used by defense lawyers to challenge children's statements. This development also sparked controversy. In a high-profile exchange in the *Cornell Law Review*, Lyon (1999) and Ceci and Friedman (2000) debated the necessity and appropriateness of the introduction of these new studies into expert testimony in child abuse cases. Two issues seemed to be at the heart of the debate. One issue, which is not addressed in the present study, was whether actual child witnesses experience the type of highly suggestive interrogations employed in the studies.⁴ The second issue, which is addressed here, concerned whether testimony about the studies is helpful to jurors. Lyon argued that the evidence gleaned from the studies is unnecessary, basing this argument on results of a few surveys of individuals' attitudes that suggest they already believe that young children are prone to false reports, and thus expert testimony unfairly stokes the skepticism of jurors who already distrust children's claims. Ceci and Friedman argued that the studies and psychological research in general provide important insights that go beyond jurors'

⁴Ceci and Friedman reported that these types of interrogative practices are sometimes employed, as indicated in dramatic examples from well-publicized cases (e.g., *McMartin Preschool*, Kelly Michaels; *Wenatchee*, Washington; *Jordan Minnesota*; *Kern County*, California). Lyon reported that these types of interrogations are rare in current interview practices with children, as demonstrated in recent studies of actual investigative interviews conducted by Sternberg, Lamb, Esplin, and Baradaran (1999) and Warren et al. (1999).

conventional wisdom regarding children's suggestibility. Because prior surveys have only included a few items relevant to children's disclosure of abuse, and no surveys have examined, in a comprehensive manner, what jurors believe concerning children's memory, suggestibility, reactions to and disclosure of abuse, it has not been possible to resolve the debate.

Legal Standards for Admissibility of Expert Testimony

The traditional standard for the admissibility of scientific evidence, first articulated in *Frye v. United States* (1923), simply required that there be a consensus in the scientific community as to the reliability of the evidence. Although the *Frye* standard is still followed in some state court systems, including those in New York and California, this standard has been supplanted in Federal courts and a number of other states by a newer standard, announced in the U.S. Supreme Court decision in *Daubert v. Merrell-Dow Pharmaceuticals* (1993). The *Daubert* standard requires the trial judge to determine first that the proffered expert testimony is "reliable," which means that it is derived from a valid scientific method, and second that it will assist the trier-of-fact.

This second requirement for admissibility depends, in part, on the knowledge and beliefs of the trier-of-fact. Thus, if, as Lyon (1999, 2002) suggested, jurors already know that which an expert could validly tell them, testimony from the expert does little to assist the trier-of-fact, and the case for admissibility of the expert is weak. According to Ceci and Friedman, however, "There is no reason to assume that the average potential juror, much less the overwhelming majority of jurors, has a good understanding of all the insights that decades of psychological research have yielded" (2000, p. 101). The only support Ceci and Friedman provided for this assertion was a request for readers to ask themselves questions about preschoolers' abilities (p. 101, note 286). Given the significance of the topic and the potential ramifications of the inclusion of helpful versus prejudicial expert testimony, scientific study of the conventional wisdom of jurors is essential.

Previous Surveys of Jurors' Knowledge

Prior studies concerning expert witnesses have largely focused on mock jurors' perceptions of the usefulness of expert testimony and the effects of expert testimony on juror decision-making (e.g., Cooper, Bennett, & Sukel, 1996; Krauss & Sales, 2001). Relatively few studies have involved child sexual abuse (e.g., Crowley, O'Callaghan, & Ball, 1994; Kovera, Gresham, Borgida, Grey, & Regan, 1997; Rudy, 1996), and researchers have rarely attempted to determine what jurors generally know about children's eyewitness capabilities and limitations. Such information is critical to the question of whether expert testimony should be admitted in court in cases involving child witnesses (but see McAuliff & Kovera, 2005). If jurors are aware of factors that research reveals influence children's memory accuracy and suggestibility, then the presentation of this research by an expert is unnecessary. If, however, jurors do not understand the range of factors that influence children's reports, knowledge of research findings may be particularly helpful.

One study of jurors' knowledge was conducted by Morison and Greene (1992), who surveyed jurors in Colorado Springs about their beliefs regarding children's experiences and disclosure of sexual abuse. Jurors appeared well informed in terms of children's general reactions to abuse (e.g., that children may be reluctant to disclose abuse, perpetrators may not use force, and children's reactions may vary), but they were less accurate with regard to the existence of physical evidence in sexual abuse cases. Only a few questions were asked about children's suggestibility and false reports, and jurors' responses suggested skepticism: More than half the jurors agreed that, "Children are easily manipulated into giving false reports about sexual abuse." Although research clearly demonstrates that children can be misled into reporting false events, including those that are negative or involve salient body contact, it does not appear "easy" to elicit such claims in children. Instead, highly suggestive interrogation practices are necessary to elicit such false reports (e.g., Leichtman & Ceci, 1994; Thompson et al., 1997).

In another study, Kovera and Borgida (1997) compared knowledge of children's reactions to and disclosure of abuse among college students, community members, and child sexual abuse experts (members of the International Society for Traumatic Stress Studies). Most community and college students were accurate in agreeing that children may retract true reports of sexual abuse and that children can distinguish between fantasy and reality. However, as in Morrison's and Greene's survey of jurors, over half (53%) of Kovera's and Borgida's respondents agreed with the statement, "Children are easily manipulated into giving false reports about sexual abuse" (only students were asked this question). This again highlights a noteworthy level of skepticism among potential jurors relative to skepticism among experts, among whom only 20% agreed. Finally, in Kovera and Borgida's survey, few differences emerged between the community and college student samples. Instead, both groups seemed to be responding quite similarly, although often differently than respondents in the expert group.

Two other surveys are indirectly relevant to the focus of the present study. In one, conducted by Corder and Whiteside (1988), North Carolina jurors were questioned about the type of evidence they felt they needed to make decisions in cases involving child victims. Most said they would need to rely on professional testimony to decide guilt or innocence in such cases. However, respondents were asked only about expert testimony concerning when children were lying, not when they were being misled, for example, by suggestive interrogation. Expert testimony on children's suggestibility may be more relevant for jurors than expert testimony on intentionally deceptive claims of sexual abuse, particularly with younger children.

The second study, conducted by Schneider (1994), compared jurors in Philadelphia with mental health professionals from the American Professional Society on the Abuse of Children and the Forensic and Clinical Psychology Associates organizations regarding their knowledge of childhood sexual abuse. Participants read statements and selected responses that they believed to be correct from multiple-choice options. Like the respondents in Kovera and Borgida's (1997) survey, participants knew that young children can differentiate between the truth and a lie, that recanting does not mean that a child was lying, and that inconsistent testimony is not necessarily false. Only a few questions in the study focused on children's

suggestibility, none of which differentiated between children's intentional lies and unintentional errors (e.g., in response to suggestive questions). A small minority of respondents (experts or jurors) was aware that children might falsely report sexual events.

Although these surveys provide important information about jurors' general knowledge of the consequences of child sexual abuse and children's disclosure of abuse, numerous questions remain. First, the surveys did not explicitly focus on jurors' knowledge of children's suggestibility. Yet, experts are often asked to provide social framework evidence specifically about scientific findings concerning children's suggestibility and false memories. Second, only two of the surveys compared potential jurors' and college students' knowledge. Although neither study found clear consistent differences between the two groups, findings from studies of juror-decision making in child sexual abuse cases have revealed that increasing juror age is associated with greater skepticism of children's reports and reduced guilt judgments (e.g., Goodman et al., 1998). Given that college student samples are younger than juror samples and that most jury decision-making research uses college student samples, it is important to assess the extent to which knowledge overlaps between the two groups. Third, the studies did not examine jurors' knowledge of specific findings from empirical research, despite the fact that this research forms the basis for experts' testimony. Fourth, the surveys were conducted in the 1980s and 1990s. Empirical research on children's suggestibility has increased exponentially in the past decade. It is not known as to whether jurors' knowledge has changed since this new scientific evidence has been collected and disseminated.

In this study, we surveyed both actual jurors and college students about their beliefs concerning children's memory abilities, suggestibility, disclosure of abuse, and reactions to abuse. We also assessed respondents' ability to infer the results of specific well-known studies of children's memory and suggestibility. We investigated whether respondents' beliefs and knowledge were related to individual-difference characteristics, such as age and gender. We anticipated that the college and jury samples would differ in age and that, with age, individuals' skepticism of children (e.g., estimates of false reports rates in studies) would increase. Finally, we conducted the study in California, a venue where jurors have not previously been surveyed about these issues. Our results provide important information about jurors' knowledge of child witnesses' abilities and limitations and offer insights directly relevant to the questions of whether expert testimony is appropriate and when such testimony is most (and least) needed.

METHOD

Participants

Participants included 169 jury-eligible undergraduate students (74 males) and 148 individuals just dismissed from jury duty at a county courthouse (76 males). In exchange for participation, college students received course extra credit and jurors received \$10.00. The average ages of the student and jury samples were 20 and 45 years, respectively.

Questionnaires

Demographic Questionnaire

This questionnaire elicited information concerning participants' age, gender, ethnicity, highest level of education (1, high school degree; 2, vocational school or community college degree; 3, 4-year college degree; and 4, graduate or professional degree), and status as a parent. It also asked how much experience participants had working with children (on a 7-point scale, none to extensive), whether they had ever been a victim of a serious crime or served on a jury, and whether they support the death penalty.

Beliefs Questionnaire

A 29-item questionnaire assessed participants' beliefs about children's memory, suggestibility, reactions to sexual abuse, and disclosure of sexual abuse. First, participants indicated their agreement with a set of 27 statements on a scale from 1 = strongly disagree to 6 = strongly agree (see Tables 2–5). Statements referred explicitly to either a 4-year-old or an 8-year-old child and were heuristically categorized into four domains: (a) *Memory and ability to testify*: e.g., “If a [4/8-year-old] child has been repeatedly and painfully sexually abused as an infant, he/she can remember it”; (b) *Suggestibility and interrogation*: e.g., “[Four/Eight-year-old] children's descriptions of events are more accurate when they are questioned in a serious manner by a law enforcement officer than when they are questioned in a supportive manner by a social worker,” “[Four/Eight-year-old] children are no more influenced by leading questions than are adults”; (c) *Reaction to sexual abuse*: e.g., “Chronic bed-wetting is almost always indicative of having been sexually abused”; and (d) *Disclosure of abuse*: e.g., “When a “[4/8-year-old] child's description of sexual abuse is disclosed over time, with more details being reported each time the child is interviewed, this indicates that the story is true.”

Second, two questions asked participants to indicate the youngest age at which they believed a child should be allowed to testify and to estimate the percentage of sexual abuse claims (made by either 4-year-olds or 8-year-olds) that are false.

Research Knowledge Questionnaire

This questionnaire assessed participants' ability to infer or guess the outcomes of several well-known studies of children's memory and suggestibility. Overviews of four studies were presented along with questions about the studies' results.

Mr. Science study (Poole & Lindsay, 1995). The first study concerned children's reports of false details regarding a visit to their class by a man who performed science demonstrations. The overview explained that, three months after a visit with Mr. Science, 3- and 8-year-old children were read a story about the visit by their parents on three consecutive nights. The story included details that had and had not occurred, including a false detail that Mr. Science put something yucky in the children's mouths. Participants were asked to estimate the percentage of 3- and 8-years-olds who falsely assented that Mr. Science put something yucky in their mouth

when interviewed a day after the story had been read the final time by their parents and again a month later, after they had been told by the interviewer that some of the things in the story had not really happened.

Mousetrap study (Ceci, Huffman, Smith, & Loftus, 1994). The second overview described a study conducted by Ceci et al. (1994) concerning children's false assents of having their hand caught in a mousetrap. The overview explained that 3–6-year-old children were questioned in a leading manner on 7–10 occasions about various true and false events (including a false event in which the child got a hand caught in a mousetrap and had to go to the hospital to have the mousetrap removed). Participants were asked to estimate the percentage of false events that, in the final interview, 3–4- and 5–6-year-olds mistakenly agreed had happened.

Janitor study (Thompson et al., 1997). The third overview described a study in which 5- and 6-year-olds reported the activities of a janitor whom they had observed cleaning some toys. The children were interviewed by a person claiming to be the janitor's boss. In the interview, the "boss" repeatedly suggested that the janitor, who had cleaned the toys, was actually playing and was, therefore, being bad. Participants were asked to estimate the percentage of children who answered all of the questions correctly about the janitor's behavior and later told their parents the truth about his behavior.

Medical checkup study (Saywitz et al., 1991). In the final overview, which described a study of girls' reports of true or false genital touch, participants were informed that 5–7-year-old girls were given a physical examination that either did or did not include a vaginal examination. A week later, the girls were interviewed about their experience. Participants were asked to estimate the percentage of girls who reported that their vaginas had been touched (correctly and incorrectly) when (a) asked a free-recall question, (b) asked to show what happened with an anatomically detailed doll, and (c) the interviewer pointed to the doll's genital area and asked whether the child had been touched there.

Procedure

Student Participants

College students were recruited from the social sciences subject pool. Following their informed consent, they were administered the demographic questionnaire, the belief questionnaire, and the research knowledge questionnaire. For half of the students (distributed randomly), the belief questionnaire referred to 4-year-olds; for half, it referred to 8-year-olds. After completing the questionnaires, students were debriefed and shown the results of the studies described in the research knowledge questionnaire.

Jury Pool Participants

Jurors in a suburban county in Southern California were recruited at the county courthouse. The rules of this court require individuals summoned for jury duty to wait in a large assembly room until they are assigned to a courtroom. Many individuals wait in the assembly room the entire day without being assigned. Those

sent to a courtroom but not seated on a jury and those who complete their service on a jury return to the assembly room at the end of the day to be released from jury service. Thus, many of the eligible jurors are in the assembly room at the end of the day and are available to be approached about research.

On two separate days, before releasing the assembled jurors, the courthouse staff informed jurors of the opportunity to participate in this study. Those who agreed to participate remained in the assembly room after being formally released from duty, and the study materials were distributed. The materials consisted of a consent form and the three study questionnaires. Again, for half of the participants, the belief questionnaire referenced 4-year-olds, and for half, the questionnaire referenced 8-year-olds. After completing the questionnaires, participants were debriefed, paid, and given copies of the results from the research studies that were described.

Although we could not assess formally the percentage of eligible individuals who took part in the study, our estimate is that approximately 90% of the potential jurors in the assembly room completed the survey. Comparisons between the ethnic characteristics of our juror sample and the county population suggest that study participants were disproportionately likely to be Caucasian and unlikely to be Hispanic (i.e., 69% of our sample was Caucasian, whereas 50% of the residents of the county are Caucasian; 10% of our sample was Hispanic, whereas 33% of county residents are Hispanic). The percentages of Asian Americans (15% in our sample and the county population), African Americans (3% in our sample, and 2% in the county population) and other/multi-ethnic respondents (3% in our sample, 2% in the county population) were comparable (<http://www.oc.ca.gov/>). It is not clear, however, what percentage of individuals across each ethnic group in the county is eligible for jury duty. Of note, our sample parallels the distribution of California voters from which many potential jurors are recruited, according to California voter registration data; in 2002, Caucasians constituted 75% and Hispanics 12% of registered voters (Reyes, 2001).

RESULTS AND DISCUSSION

Results are presented and discussed in the following order: First, differences in demographic characteristics of the two samples; second, results of analyses investigating respondents' agreement with statements regarding children's memory, suggestibility, and reactions to and disclosure of abuse; third, individual differences in the accuracy of participants' beliefs; fourth, participants' opinions regarding children's testimony and false reports of abuse; and fifth, respondents' estimates of the results of specific studies of children's memory and suggestibility. For inferential statistics reported, estimates of the effect sizes (Cohen's d for t -tests, and partial η^2 , denoted η^2 , for ANOVAs) are included.

Demographic Differences Between College Student and Jury Samples

Demographic characteristics of the college-student and jury-pool samples are presented in Table 1. As expected, the jury pool sample was significantly older than

Table 1. Demographic Characteristics of College Students and Jury Pool Members

	College students ($N = 169$)	Jury pool ($N = 148$)
Mean age (years)	20 (3)	45 (13)
Ethnicity (%)		
Asian	53	15
Caucasian	29	69
Hispanic	11	10
African American	0	3
Multi-ethnic or other	7	3
Education (%)		
Completed high school	—	10
Completed vocational school or 2 years of college	—	37
Completed college (bachelor's degree)	—	34
Completed post graduate training	—	19
Percent who have children	2	33

the college sample, $t(314) = 25.10$, $p < .001$. The ethnic distribution also varied significantly, $X^2(4) = 64.80$, $p < .001$: A higher percentage of the college student sample was Asian American, and a higher percentage of the jury pool sample was Caucasian. Also, whereas only 2% of the college students had children, 66% of the jury pool participants had children. We expected that that the juror and college samples' beliefs might vary (with the jury sample being more skeptical of children) based on former studies suggesting increasing age is associated with greater skepticism. As reported below, few group differences emerged in participants' beliefs. Moreover, findings were identical when age was statistically controlled in analyses. For ease in interpretation, analyses without controlling for age are reported.

Knowledge of Memory, Suggestibility, and Child Sexual Abuse

For judges who make decisions about the admissibility of expert testimony and scientists who study juries, the *extent* and *accuracy* of jurors' knowledge are key issues. Accordingly, we examined what participants think they know about children's memory ability and suggestibility, children's reactions to abuse and to interrogation, and professionals' ability to detect abuse. We also assessed whether participants' beliefs about these issues are consistent with findings from empirical research that might be presented by an expert. Because there were no differences between the college student and juror samples (details discussed shortly), here we report findings collapsed across groups.

Table 2 displays participants' responses to each of the 27 statements in the beliefs questionnaire. The table includes: (a) the overall mean score for the sample on the 6-point scale of agreement (responses ranged from 1 = strongly disagree to 6 = strongly agree); (b) the percentage of participants who agreed with the statement (collapsed across three responses: 4, slightly agree; 5, moderately agree; or 6, strongly agree); (c) the percentage of respondents who answered, "do not know"

Table 2. Mean Agreement with Statements on the Belief Questionnaire from the Memory and Ability to Testify Category

Belief statement	Mean score of sample	Who agreed (%) (ratings 4–6)	Who were correct (%)	Who reported do not know (%)
Children can remember repeated, common experiences but not experiences that happen just once. ^a	2.61	24	70	6
Children cannot remember events well enough to be reliable witnesses in court. ^{a,e}	2.89	31	66	3
Children can distinguish between the truth and a lie. ^{b,d}	4.23	70	70	4
If a child has been repeatedly and painfully sexually abused as an infant, he/she can remember it. ^{a,d}	3.69	48	36	16

Note. Participants rated their agreement with each statement on a scale from 1 = strongly disagree to 6 = strongly agree; a do-not-know option was also listed. Participants who agreed with true statements and disagreed with false statements were classified as correct; statements were heuristically classified as true or false based on empirical research, case studies, and former expert testimony. To facilitate comparisons across Tables 2–5, the same superscript notations are used.

^aStatements classified as false, and disagreement responses (scores of 1–3) were considered accurate.

^bStatements classified as true and agreement responses (scores of 4–6) were considered accurate.

^cStatements that are not included in the overall composite accuracy score.

^dStatements that reflect trust in children’s abilities.

^eStatements that reflect skepticism about children.

to each statement; and (d) the percentage of participants whose response was “correct.” To create this last score, we classified each of the statements as either true or false based on information collected from several sources: empirical research, expert testimony provided in legal cases, and conclusions that many experts have drawn from the literature. Responses were considered correct if participants disagreed with a false statement (gave a rating of 1, 2, or 3) or agreed with a true statement (gave a rating of 4, 5, or 6).

Although we expect that most experts in the field would agree with most of our classifications, some classifications are less fully agreed upon than others. For example, the statement that children can remember highly distressing events from infancy was classified as false because research has consistently failed to support such a possibility (e.g., Peterson, 1999; Quas et al., 1999; see Howe, Courage, & Edison, 2003). In this case, our classification would likely be met with minimal controversy. In contrast, there might more disagreement about the statements, “Most children can be manipulated into making a false claim about sexual abuse” and “Repeatedly asking children specific questions, such as ‘Did he touch your private parts,’ often leads to false reports of sexual abuse,” both of which we classified as true. Empirical research has obviously not directly tested the veracity of these statements.

Additionally, although researchers generally agree that certain types of highly leading and coercive interview tactics lead to dramatic errors in children's reports, they are less likely to agree about whether yes/no questions constitute leading questions (e.g., in some studies, such questions are labeled "leading;" e.g., Scullin & Ceci, 2001; in other studies, such questions are labeled "specific" and contrasted with "misleading" questions; e.g., Goodman, Bottoms, Schwartz-Kenney, & Rudy, 1991; Quas & Schaaf, 2002). Researchers also continue to debate whether repeated questions in and of themselves increase errors, or whether the use of highly leading questions rather than question repetition leads to children's inaccuracies (e.g., Bruck et al., 2002; Quas, Malloy, Goodman, Melinder, Schaaf, & D'Mello, 2005). When clear empirical evidence was not available as a basis for classifying statements, and when the veracity of statements remained somewhat controversial, we based our classifications on facts of actual legal cases (e.g., we assume that children can be led to falsely claim sexual abuse based on cases in which such claims were shown to have been highly improbable) and our belief that the majority of experts would endorse the classification. Of course, the latter interpretation is an empirical question worthy of direct study on its own, and we return to this issue again in the Discussion.

Do People Think They Know

Results for the "do-not-know" response (last columns in Tables 2–5) reveal that most respondents believed that they knew the answers to most questions. For 22 of the 27 statements, 88% or more of the participants gave a response rather than claiming they did not know. The statement to which the fewest respondents (3%) answered "do not know," which fell in the *Memory and ability to testify* category, was, "Children cannot remember events well enough to be reliable witnesses in court" (Table 2). Only two statements were given "do not know" responses by more than 20% of the participants. Both statements were in the *Reactions to sexual abuse* category: "Most children who have been sexually abused display bizarre sexualized behavior" (26% did not know), and "Chronic bedwetting by a child is almost always indicative of having been sexually abused" (22% did not know, see Table 4). Thus, participants were somewhat more likely to believe that they knew about children's memory and suggestibility than about children's reactions to sexual abuse. In other words, they were more willing to agree or disagree with the former types of statements and to say, "do not know" to the latter type of statements.

What People Believe

The third column in Tables 2–5 shows the percentage of participants whose response to each belief statement was judged to be correct—i.e., the percentage who agreed with true statements and disagreed with false statements.

Memory and ability to testify. Four statements were classified as tapping respondents' general beliefs about children's memory and ability to testify (see Table 2). Respondents' answers to most of these statements indicated that their beliefs are

Table 3. Mean Agreement with Statements on the Belief Questionnaire from the Suggestibility and Interrogation Category

Belief statement	Mean score of sample	Who agreed (%) (ratings 4-6)	Who were correct (%)	Who reported do not know (%)
Children are more suggestible about events that they do not understand than about events that they understand. ^b	4.46	72	72	15
Children's descriptions of events are more accurate when they are questioned in a serious manner by a law enforcement officer than when they are questioned in a supportive manner by a social worker. ^a	2.67	27	64	9
Children are no more influenced by leading questions than are adults. ^{a,c}	2.97	38	55	7
Repeatedly asking children specific questions, such as, "Did he touch your private parts?" often leads them into making false claims of sexual abuse. ^{b,d}	3.81	58	58	7
Repeatedly asking children general open-ended questions, such as "What happened? What else happened?" often leads them into making false claims of sexual abuse. ^{a,d}	3.30	46	47	7
Asking children to use anatomically detailed dolls (dolls with genitalia) to show "what bad things happened" is likely to lead to false reports of sexual abuse. ^{b,d,e}	3.07	30	30	11
Children are sometimes led by an adult into reporting that they have been sexually abused when they have not. ^{b,d}	4.23	71	71	10
A child cannot describe sexual abuse unless he/she actually experienced it. ^{a,c}	3.96	56	36	8
Children sometimes make up stories about having been sexually abused when they actually have not. ^{b,d}	3.13	38	38	11
Children sometimes come to believe that they were sexually abused when they really were not. ^{b,d}	3.40	46	46	12
Most children can be manipulated into making a false claim about sexual abuse. ^{b,d,e}	4.18	70	70	6

Table 3. Continued

Belief statement	Mean score of sample	Who agreed (%) (ratings 4-6)	Who were correct (%)	Who reported do not know (%)
Asking a child to name body parts on an anatomically-detailed doll (a doll with genitalia) is likely to lead to false reports of sexual abuse. ^{a,d}	2.94	26	54	20
A psychologist can tell whether a child's description of an event has been influenced by another adult. ^a	3.98	68	25	7

Note. Participants rated their agreement with each statement on a scale from 1 = strongly disagree to 6 = strongly agree; a do-not-know option was also listed. Participants who agreed with true statements and disagreed with false statements were classified as correct; statements were heuristically classified as true or false based on empirical research, case studies, and former expert testimony. To facilitate comparisons across Tables 2-5, the same superscript notations are used.

^aStatements classified as false, and disagreement responses (scores of 1-3) were considered accurate.

^bStatements classified as true and agreement responses (scores of 4-6) were considered accurate.

^cStatements that reflect trust in children's abilities.

^dStatements that reflect skepticism about children.

^eStatements that are not included in the overall composite accuracy score.

Table 4. Mean Agreement with Statements on the Belief Questionnaire from the Reactions to Sexual Abuse Category

Belief statement	Mean score of sample	Who agreed (%) (ratings 4-6)	Who were correct (%)	Who reported do not know (%)
Most children who have been sexually abused display bizarre sexualized behavior. ^a	3.54	41	33	26
Chronic bedwetting by a child is almost always indicative of having been sexually abused. ^a	1.99	9	69	22
A physical examination by a doctor will almost always show whether or not a child has been sexually abused. ^a	3.92	57	33	10

Note. Participants rated their agreement with each statement on a scale from 1 = strongly disagree to 6 = strongly agree; a do-not-know option was also listed. Participants who agreed with true statements and disagreed with false statements were classified as correct; statements were heuristically classified as true or false based on empirical research, case studies, and former expert testimony. To facilitate comparisons across Tables 2-5, the same superscript notations are used.

^aStatements classified as false, and disagreement responses (scores of 1-3) were considered accurate.

^bStatements classified as true and agreement responses (scores of 4-6) were considered accurate.

^cStatements that are not included in the overall composite accuracy score.

^dStatements that reflect trust in children's abilities.

^eStatements that reflect skepticism about children.

Table 5. Mean Agreement with Statements on the Belief Questionnaire from the Disclosure of Abuse Category

Belief statement	Mean score of sample	Who agreed (%) (ratings 4–6)	Who were correct (%)	Who reported do not know (%)
When a child's description of sexual abuse is disclosed over time, with more details being reported each time the child is interviewed, this clearly indicates that the child's description is false. ^{a,e}	2.74	25	67	8
Inconsistencies in a child's report of sexual abuse indicate that the report is false. ^{a,e}	2.96	29	67	5
Most children who are sexually abused tell someone right away. ^a	2.00	9	84	7
Children who retract (take back) their stories about sexual abuse were probably lying in the first place. ^{a,e}	2.76	21	69	10
Children who have been sexually abused will not deny it if asked by a trusted adult. ^a	3.68	51	43	6
Children sometimes make up false claims of sexual abuse to get back at an adult. ^{a,e}	2.86	33	57	10
When a child's description of sexual abuse is disclosed over time, with more details being reported each time the child is interviewed, this indicates that the child's description is true. ^{a,d}	3.74	52	36	12

Note. Participants rated their agreement with each statement on a scale from 1 = strongly disagree to 6 = strongly agree; a do-not-know option was also listed. Participants who agreed with true statements and disagreed with false statements were classified as correct; statements were heuristically classified as true or false based on empirical research, case studies, and former expert testimony. To facilitate comparisons across Tables 2–5, the same superscript notations are used.

^aStatements classified as false, and disagreement responses (scores of 1–3) were considered accurate.

^bStatements classified as true and agreement responses (scores of 4–6) were considered accurate.

^cStatements that are not included in the overall composite accuracy score.

^dStatements that reflect trust in children's abilities.

^eStatements that reflect skepticism about children.

compatible with scientific consensus. For instance, 70% believe that children can remember events that occurred once as well as events that are repeated, a pattern consistently supported by empirical research (e.g., Farrar & Goodman, 1992; Goodman, Quas, Batterman-Faunce, Riddlesberger, & Kuhn, 1997; Hudson, Fivush, & Kuebli, 1992; Lucariello & Nelson, 1985). Slightly fewer, but nonetheless a majority (66%) of participants believe that children can remember events well enough to be reliable witnesses in court. Although children's memory reports are often incomplete, research indicates that even children as young as 2 or 3 years of age are able to talk about past events (e.g., Fivush & Schwarzmüller, 1998; Peterson, 1999) and, when no external suggestions have been provided, children's narrative reports are often

highly accurate, with the amount of information they provide about their experiences steadily increasing across the preschool and early school age years (e.g., see Kail, 1990; Ornstein & Haden, 2002). Seventy percent of the participants also correctly agreed that, "Children can distinguish between the truth and a lie." Research has in fact revealed that even young children can distinguish between the truth and a lie, especially when questions are asked in a concrete, developmentally appropriate manner (Bussey, 1992; Haugaard, 1993; Lyon & Saywitz, 1999). In Schneider's (1994) survey, 58% of the jurors agreed with the statement that children aged 2–8 years know the difference between the truth and a lie. The smaller percentage of respondents who agreed with the statement in Schneider's study compared with our study may be a result of the fact that the statement they were given included 2-year-olds, whereas our study only asked about 4- or 8-year-olds.

Although the majority of participants responded correctly to the first three statements about children's memory, only 36% correctly disagreed with the statement, "If a child has been repeatedly and painfully sexually abused as an infant, he/she can remember it." Research clearly indicates that events which took place before the age of 3 years are unlikely to be remembered later in childhood or adulthood, a phenomenon termed infantile amnesia (e.g., Quas, Goodman, Bidrose, Pipe, Craw, & Ablin, 1999; Usher & Neisser, 1993; West & Bauer, 1999). Forty-eight percent of our respondents think 4- or 8-year-old children can remember such events, and 16% said they did not know. That so many people hold false beliefs about children's ability to remember events from infancy could be especially problematic in a legal case involving alleged abuse that took place before the offset of infantile amnesia. An expert's testimony on infantile amnesia could be critical to such a case.

Suggestibility and interrogation. The largest number of statements fell in the second category, which tapped respondents' beliefs regarding children's reactions to and capabilities in forensic interview situations. Participants' mean response, the percent of respondents who agreed, the accuracy of respondents' perceptions, and the percent of respondents who answered "do not know" are presented in Table 3. As can be seen, the majority of participants recognized that suggestibility could lead to inaccuracies in children's reports, including false reports of sexual abuse: 72% correctly agreed that children are more suggestible about events that they do not understand than events that they understand; 71% agreed that children are sometimes led by an adult into reporting that they have been sexually abused when they have not; and 70% agreed that children can be manipulated into reporting falsely that they have been sexually abused.

Respondents' beliefs are consistent with a large body of scientific research (Bruck et al., 2002; Ceci et al., 1994; Garven et al., 1998; Thompson et al., 1997) and highly publicized legal cases (e.g., see Ceci & Bruck, 1995) which reveal that, under extreme circumstances, children will claim that sexual abuse occurred, even when its occurrence would have been virtually impossible. Our findings thus support Lyon's (1999, 2002) view that jurors generally appreciate (without being told by experts) that children can be suggestible. It is important to note, however, that belief in children's suggestibility was not universal. A sizeable minority (incorrectly) disagreed with these statements about children's suggestibility.

Fewer participants but still a majority (64%) correctly disagreed with the statement that serious questions by a law enforcement officer are more likely to elicit accurate descriptions than supportive questions by a social worker. Research indicates that, as the majority of respondents believe, children are more accurate and less suggestible when questioned by a supportive interviewer (e.g., Carter, Levine, & Bottoms, 1996; Davis & Bottoms, 2002). A majority (58%) also correctly agreed that repeating direct questions (e.g., Did he touch your private parts?) could lead to false reports of sexual abuse. Although there is some controversy regarding the ease with which direct or repeated questions elicit abuse reports in nonabused children (e.g., Goodman et al., 1991; Bruck et al., 1997), errors in children's reports of other events increase following repeated, leading questions (e.g., Leichtman & Ceci, 1994; Thompson et al., 1997), and direct questions are considered leading by many researchers (e.g., Scullin & Ceci, 2001; Poole & Lamb, 1998; Sternberg, Lamb, Esplin, Orbach, & Herschkowitz, 2002). Just over half of the participants (54%) also believe, correctly, that asking children to name body parts on an anatomical doll will not necessarily lead to false reports of sexual abuse. Empirical research indicates that leading or suppositional questions need to accompany the presentation of anatomical dolls to elicit false reports (e.g., Bruck et al., 1997; Saywitz et al., 1991). Again, however, despite accurate responses among a majority of respondents, a large minority demonstrated inaccurate beliefs.

Slightly over half of the participants (55%) correctly disagreed with the statement, "Children are not more influenced by leading questions than are adults." In Kovera and Borgida's (1997) survey a somewhat larger number of respondents (64% of experts, 75% of students, and 76% of the community sample) correctly disagreed with this statement. In both studies, however, a sizeable number of individuals was inaccurate. Age differences in memory, suggestibility, and false reports are arguably the most consistent findings in the field. With very few exceptions, memory improves with age and suggestibility decreases (see Ceci & Bruck, 1993, 1998). Given the robustness of age differences, it is surprising that so many participants (38%) appear to believe, incorrectly, that children are no more influenced by leading questions than adults. Of course, some respondents may have incorrectly agreed with the statement not because they believe that children are resistant to leading questions but because they know that adults are also susceptible to suggestion (e.g., Loftus & Pickrell, 1995; see Loftus, 2003). Even so, their responses reflect a failure to recognize age differences on this dimension.

Most respondents (68%) agreed, incorrectly, with the statement: "A psychologist can tell whether a child's description of an event has been influenced by another adult." Research has yet to examine whether professionals can distinguish true and false allegations when they view elaborate, lengthy interviews with children. However, Leichtman and Ceci (1994) showed brief videos of children incorrectly answering leading questions during mock forensic interviews to researchers and clinicians who work in the area of children's testimony. The professionals often rated children's false reports as true, suggesting a lack of detection abilities on the part of experts. Corder and Whiteside (1988) found similarly high levels of perhaps unwarranted trust in professionals' ability to detect abuse. Most jurors in their study believed that mental health professionals can tell whether a 3–5-year-old is lying about

sexual abuse and whether a child has been sexually abused, even when no physical evidence is available. Corder and Whiteside's respondents also said that they would rely on mental health professionals to help them evaluate children's statements. Thus, both our and Corder and Whiteside's results suggest that individuals trust professional evaluations to a greater extent than is warranted according to research. Of course, respondents to both surveys may have been basing their judgments on beliefs about professionals' abilities when they evaluate children during lengthy assessments or interviews rather than, as in Leichtman and Ceci's (1994) study, when they evaluate children's veracity only after listening to brief interviews. Nonetheless, because empirical studies have yet to examine professionals' detection abilities following exposure to elaborate, lengthy interviews with children, jurors may well need to be educated, for instance, by an expert, regarding limitations in or at least the lack of available evidence concerning professionals' diagnostic abilities.

Only 46% of the participants correctly agreed that children sometimes come to believe that they were sexually abused when they were not. Although researchers have not been able to examine children's false beliefs about sexual abuse directly, research indicates that children can internalize suggested stories and come to believe that suggested events really occurred (Ceci et al., 1994; but see Huffman, Crossman, & Ceci, 1997). Only 47% of the participants correctly disagreed that repeatedly asking children general open-ended questions such as "What happened? What else happened?" leads to false claims of sexual abuse. Research, however, indicates that open-ended questions, per se, do not create false reports in children (e.g., Brown & Pipe, 2003; Quas & Schaaf, 2002). Thus, jurors may interpret nonleading questions as suggestive and may need assistance from experts regarding not only what constitutes a suggestive interview, but also, and of importance, what does not.

Even fewer respondents (30%) correctly agreed that having children use anatomically detailed dolls to show "what bad things happened" is likely to lead to false reports of sexual abuse. We judged this belief statement to be true based on results of research by Bruck, Ceci, Francoeur, & Renick (1995), which revealed that preschool-age children will err when asked to use a doll and props to demonstrate what "could" have happened during a particular event. However, although questions like those used by Bruck and her colleagues produced incorrect responses, other studies (employing different forms of interrogation) indicate that a majority of children can be accurate when asked to demonstrate past events with dolls and props, even when they are asked falsely leading questions about genital touching (e.g., Goodman et al., 1991; Steward & Steward, 1996). Accordingly, whether dolls and props promote accurate or inaccurate accounts appears to depend on the details of how questions are phrased. Thus, that a majority of our respondents erred on this particular question is not surprising given the academic controversy that exists over the precise conditions under which such dolls are suggestive. More generally however, participants' responses across the two questions suggest that the majority does not believe the view that dolls can promote false claims of sexual abuse, a possibility supported by some scientific researchers (e.g., Bruck et al., 1995).

A final limitation in participants' knowledge of children's suggestibility was evident in their responses to the statement, "A child cannot describe sexual abuse unless he or she has actually experienced it." Only 44% correctly disagreed with

this statement. Schneider (1994) similarly found that a minority of jurors (30%) correctly agreed that a nonabused child can describe sexual abuse. The incorrect majority view in both studies is consistent with the practice of using precocious sexual knowledge as an indicator of a history of sexual abuse. However, studies of children's knowledge of sexual matters reveal that, by age 4, children have at least a limited understanding of some sexually relevant information (e.g., they can often name some sexual body parts), and sexually abused and nonabused children's basic sexual knowledge often does not differ (Gordon, Schroeder, & Abrams, 1990a, 1990b). Thus, in theory, children might be able to describe at least some sexual activities without having been exposed to sexual abuse.

Reactions to child sexual abuse. Participants' responses to statements in the third category, which included three statements, also reveal both accuracies and inaccuracies. As is evident in Table 4, the majority (69%) of participants accurately agreed that chronic bedwetting is not necessarily indicative of sexual abuse. However, only a minority (33%) knew that most children who have been sexually abused do not display bizarre sexualized behavior. In general, no single behavior, symptom, or medical diagnosis is evident in all abused children; the prevalence rate for sexually inappropriate behavior is quite low; and some children display no overt symptoms at all (Conte & Shuerman, 1987; Gomes-Schwartz, Horowitz, & Sauzier, 1985; Kendall-Tackett, Williams, & Finkelhor, 1993; Mannarino & Cohen, 1986). Similar results were reported by Schneider (1994), who found that only a minority of jurors knew that most sexually abused children do not display sexualized behavior.

Our respondents were also ignorant about the existence of physical evidence of abuse. In particular, only 33% knew that a physical examination does not typically show whether a child has been sexually abused. Similar results were reported by Morison and Greene (1992) and Kovera and Borgida (1997): The majority of their juror, community member, and student samples believed that physical evidence exists in most sexual abuse cases and that medical experts can detect abuse. Empirical evidence, however, suggests the contrary. Medical exams often produce inconclusive results (Herbert, 1987; Mian, Wehrspann, Kaljner-Diamond, LeBaron, Winder, 1986), and physical evidence of abuse is uncommon (e.g., DeJong & Rose, 1991).

Disclosure of abuse. Participants' responses to the seven statements in the final category are displayed in Table 5. As is evident in the percentages listed in the third column, participants tended to be quite accurate. Most participants knew that children who are sexually abused may not tell someone right away (84%); that increasingly detailed disclosures of abuse over time do not clearly indicate that the abuse report is false (67%); that inconsistencies in children's claims of sexual abuse do not indicate that their claims are false (67%); and that retracting claims of sexual abuse does not mean that children were lying in the first place (69%). The majority of participants in Kovera and Borgida's (1997) survey also agreed that delays in reporting sexual abuse are common and that retraction of a report is not an indicator that children were lying in the first place. Such beliefs are consistent with empirical research. Children's true disclosures of negative experiences may occur over time and after a delay (e.g., Sorensen & Snow, 1986), and children's true reports often contain inconsistencies (e.g., Fivush & Schwarzmuller, 1995). The majority of respondents in our study (57%) also agreed that, "Children do not make up false

claims of sexual abuse to get back at an adult.” Although research has not examined whether children intentionally make false reports of abuse out of malice or spite, we think most experts would doubt the cognitive capacity of children in the age range targeted in the study (4 and 8 years) to engage in such behavior.

Two important inaccuracies in this final category, however, have direct implications for the need for expert testimony. First, only 43% of our respondents correctly agreed that, “Children who have been sexually abused may deny it if asked by a trusted adult.” Research indicates that children often deny having been sexually abused, even when clear medical evidence, such as a sexually transmitted disease, exists (Hershkowitz et al., 2004). Our participants apparently believe that children will be more forthcoming than they often are in disclosing sexual abuse. Second, over half (52%) of the respondents incorrectly agreed that, “When a child discloses sexual abuse in increasing detail over time, with more details being reported each time the child is interviewed, this indicates that the description is true.” Inconsistencies and disclosure over time are not indicative of accuracy, and increasing detail may well be the result of suggestion rather than improved memory.

Summary. Our descriptive analysis reveals both accuracies and inaccuracies in participants’ knowledge. There were no statements that were responded to correctly by *all* participants. Yet, respondents demonstrated considerable accuracy in (a) their understanding of children’s general memory abilities (e.g., that children can recount single and repeated events accurately and can remember events well enough to be reliable witnesses), (b) their beliefs about child suggestibility (e.g., that children can be led by an adult into reporting entirely false events, including alleged sexual abuse), and (c) their understanding of how child sexual abuse is disclosed (e.g., that disclosures may be delayed, retracted and inconsistent over time). At least two-thirds of the participants gave responses indicating that they held accurate beliefs about these issues. At the same time, a substantial number of participants (more than half) lacked accurate knowledge regarding (a) infantile amnesia—children’s inability to remember events from infancy, (b) which interview tactics are more or less likely to lead to errors (e.g., direct or leading questions vs. open-ended questions, asking children to demonstrate “bad things” that happened with dolls vs. neutral naming of body parts with the assistance of dolls), (c) limitations in professionals’ abilities to detect abuse and memory inaccuracies, (d) symptoms of child abuse (e.g., that most abused children do not display bizarre sexualized behavior), and (e) the fact that children may actually believe the false events they describe and can describe and make up stories about sexual acts without having experienced them. One function of an expert witness might be to reduce the variability in jurors’ knowledge of all these issues and to provide empirically based information about the latter issues in particular.

Individual Differences in Accuracy of Beliefs

Are certain groups of individuals especially accurate in their knowledge? Are respondents more accurate about older or younger children? Do jurors’ and college students’ differ in their accuracy? To address these questions, we created a composite score ($\alpha = .73$) by averaging participants’ agreement ratings on 25 of the 27

statements from the beliefs questionnaire (correct *disagreement* items were reverse scored so that higher values reflected greater accuracy). Two items, both from the *Suggestibility and interrogation* category, were eliminated because they reduced reliability of the composite; these items are noted in Table 3.

The composite was entered into a 2 (child age: 4 vs. 8 years) \times 2 (sample: college student vs. jury pool member) \times 2 (participant gender) between-subject analysis of variance (ANOVA). The main effect of gender was significant, $F(1, 309) = 6.87$, $p < .01$, $\eta^2 = .02$, and revealed that females, $M = 3.96$, had significantly higher accuracy scores than males, $M = 3.83$. No other significant differences emerged. Of note, the sample effect was nonsignificant, suggesting that college and jury participants' beliefs were similar.

Next, participants' ethnicity, age, education, status as parents, nonparental experience with children, child and adult victimization experiences, jury service, and support for the death penalty were examined in relation to their accuracy composite score. For continuous variables, Pearson correlations were computed, respectively. For categorical variables, t-tests or ANOVAs were conducted. A significant group effect emerged when non-Hispanic Caucasians, Asian Americans, and Hispanic Americans were compared, $F(2, 293) = 18.32$, $p < .001$, $\eta^2 = .11$. (African Americans and other/multi-ethnic individuals were not included because there were too few participants in these categories.) Scheffe's post hoc comparisons revealed that Caucasian individuals, $M = 4.04$, were significantly more accurate than Asian Americans, $M = 3.70$, and Hispanic Americans, $M = 3.78$, $ps \leq .01$; the latter two groups did not differ significantly from each other. Higher level of education was associated with greater accuracy, $r(317) = .16$, $p < .01$. Finally, individuals who reported that they did not support the death penalty, $M = 4.04$ were more accurate in their overall knowledge than individuals who supported the death penalty, $M = 3.88$, $t(234) = 2.26$, $p < .05$, $d = .32$. Accuracy was not related to participants' age, status as parents, nonparental experience with children, jury service, or victimization experiences.

Although Kovera and Borgida (1997) did not examine the relations between demographic characteristics of respondents and their survey answers, both Morison and Green (1992) and Schneider (1994) investigated how basic demographic characteristics (e.g., education, age, gender) related to participant responses. Morison and Green reported that increasing education and age were associated with greater accuracy in response to some but not all statements.

With regard to gender, in Morison and Green's (1992) survey, as in ours, women demonstrated greater accuracy than men. Juror gender differences are also routinely found in studies of mock juror decision-making in child sexual abuse cases. Specifically, men are more skeptical than women about children's abuse claims and tend to rate children as less believable, credible, and competent; women tend to express more favorable attitudes toward child victims and more negative attitudes towards defendants (e.g., Bottoms & Goodman, 1994; Kovera et al., 1994; Quas, Bottoms, Haegerich, & Nysse-Carris, 2002; Schutte, 1997). Given these results, we suspected that at least some of the gender differences we observed were a function of statements tapping trust versus skepticism about children. That is, it was possible that women's greater accuracy in our survey (and possibly Morison's and Greene's)

was due to women's greater endorsement of true statements about children's capabilities and stronger disagreement with untrue statements about children's limitations. To test such a possibility, we identified a subset of statements on the belief questionnaire that reflected either trust or skepticism of child witnesses. Five statements were classified as reflecting "trust" in children (e.g., "Children can distinguish between the truth and a lie," see Notes for Tables 2 through 5 for the exact items). When participants' agreement ratings were averaged across these items and entered into a 2 (child age) \times 2 (sample) \times 2 (participant gender) ANOVA, no significant effects emerged. Nor were any significant effects evident when analyses were conducted separately for composites including only trust statements classified as true and false separately.

Thirteen statements were classified as reflecting skepticism (e.g., "When a child's description of sexual abuse is disclosed over time, with more details being reported each time the child is interviewed, this clearly indicates that the child's description is false"). These items are noted on Tables 2 through 5. Participants' average agreement with these items was examined in a 2 (child age) \times 2 (sample) \times 2 (participant gender) ANOVA. The main effect of gender was significant, $F(1, 309) = 34.31$, $p < .001$, $\eta^2 = .10$. Collapsed across correct and incorrect statements, men agreed with skeptical statements more strongly than did women (mean agreement = 3.52 for men, 3.03 for women). When agreements with correct versus incorrect skeptical statements were analyzed separately, men were more accurate for correct skeptical statements, $M_s = 3.87$ for men and 3.42 for women, and women were more accurate for incorrect skeptical statements, $M_s = 3.33$ for men and 2.67 for women, $F_s(1, 308) > 17.86$, $p < .001$, $\eta^2 \leq .06$. Thus, the gender differences in accuracy were largely due to men's and women's differential endorsement of items highlighting limitations in children's abilities. When items describing true capabilities of children were considered, gender differences did not emerge.

Similar analyses were conducted to determine whether the associations between participants' ethnicity and education and their accuracy varied depending on whether statements suggested trust or skepticism of children's reports. With regard to ethnicity, one-way ANOVAs comparing Caucasians, Hispanics, and Asian Americans were conducted with participants' overall trust and skepticism scores (regardless of the actual accuracy of the ratings) as separate dependent measures. The model was significant when participants' trust in children was considered, $F(2, 293) = 4.42$, $p = .01$, $\eta^2 = .03$. Scheffe post hoc tests revealed that Asian Americans' ratings, $M = 3.92$, were significantly higher, suggestive of their being more trusting of the veracity of children's claims of sexual abuse, than were the ratings of Caucasians, $M = 3.61$. Hispanic Americans, $M = 3.74$, did not differ from the other two groups. All but one of the items included in the trust composite were classified as false such that participants' agreement ratings (i.e., ratings that fell between 4 and 6) were considered incorrect. When only these items were examined, no significant ethnic differences emerged. When the one true statement was analyzed ("Children can distinguish between the truth and a lie"), a small but statistically significant ethnicity effect, $F(2, 281) = 3.54$, $p < .05$, $\eta^2 = .02$, emerged, with Caucasians, $M = 4.44$ reporting somewhat stronger agreement than Asian Americans, $M = 3.98$. The mean for Hispanic Americans, 4.06, fell between the other two.

Because of these seemingly opposing patterns of findings: Asian Americans appearing more trusting on one measure and Caucasian Americans appearing more trusting on the other, and the small effect sizes associated with the significant differences, we hesitate to draw inferences about the implications of the findings for jury selection or the use of experts. When participants' level of education was correlated with their trust and skepticism scores, no significant relations emerged. Thus, the significant relation between education and overall accuracy was not being driven *per se* by specific patterns of trust or skepticism on the part of our respondents.

In summary, two sets of noteworthy patterns of results emerged from the individual difference analyses. First, no significant differences emerged in overall accuracy between the college student and jury samples. This pattern was evident when participants' overall accuracy score was considered and when only those statements endorsing high trust versus skepticism in children's reports were considered. Kovera and Borgida (1997) also found relatively few differences between their college student and community sample's perceptions of children's reactions to and disclosure of child sexual abuse, and in a meta-analysis, Bornstein (1999) reported that college student and jury samples were largely similar in their judgments across a range of types of jury-decision making studies. Insofar as researchers continue to investigate laypersons' knowledge of child sexual abuse and suggestibility, including the need for expert witnesses in cases involving allegations of child sexual abuse, our results suggest that data collected from college student samples are informative and applicable to juror populations generally.

Second, a few potentially important individual-difference factors emerged as predictors of participant accuracy. Women were more accurate than men when the overall accuracy composite score was considered. Follow-up analyses revealed that women's increased accuracy seemed to be driven by men's greater skepticism in children's allegations, skepticism that was often not supported by empirical research. Ethnicity and education were also related to participants' accuracy. The findings concerning ethnicity were mixed, with Caucasians reporting less and more trust in children than Asian Americans across different items. Greater education was associated with more accurate knowledge. Because the correlations between participants' education and the averages of their ratings to the subset of items endorsing trust and skepticism in children were nonsignificant, it was not the case that educated individuals differed from less educated individuals in their trust of children's claims or skepticism about child sexual abuse allegations. Instead, perhaps with greater education, individuals become more discriminating in their judgments and increasingly recognize the complexity of evaluating children's reports of and reactions to sexual abuse. Education may also lead to greater general knowledge regarding capabilities and limitations in children's cognitive and social development, knowledge that could help individuals evaluate the survey items with increased accuracy. To the extent that these findings are replicated across settings, they suggest that certain individuals may be more prone to believe or not believe children's reports, and that certain constellation of juries may benefit more than others from information provided by an expert witness. For instance, when juries are male dominated, expert witnesses may be particularly useful in terms of overcoming males' increased tendency to be unduly skeptical of children's claims. Expert witnesses may

also benefit less educated jurors in terms of providing them with additional knowledge about both children's memory capabilities and children's susceptibility to false suggestions.

Opinions About Child Testimony and Abuse Allegations

Two other items were included in the belief questionnaire. One asked about the youngest age at which a child should be allowed to testify in a legal case. Participants' responses ranged from 1 to 21 years, with a mean age of 9.5 years. This surprisingly high mean age indicates that many jurors are skeptical about young children's testimony, although the substantial range indicates that there is no consensus. Only 6% of the respondents indicated that a child age 4 years or younger should testify; 30% said the youngest age should be 5–7 years; 32% said 8–10 years; 22% said 11–13 years; and 11% said age 14 or older. No significant differences emerged when participants' responses were related to individual difference characteristics, with one exception. Caucasians, $M = 8.47$, reported a significantly lower age than did Asian Americans, $M = 10.53$ (Hispanic Americans, $M = 10.03$, did not differ from the other two ethnicities, Scheffe's $p < .05$), overall $F(2, 275) = 10.47$, $\eta^2 = .07$. When Corder and Whiteside (1988) asked, "At what age is a child's testimony credible?" 70% of the respondents thought that children age 3 years and older could give adequate testimony about sexual abuse. Perhaps the increased media exposure to high profile cases involving false allegations of sexual abuse has led to greater skepticism regarding testimony from young children since Corder and Whiteside conducted their study.

The other question asked participants to estimate the percentage of abuse claims made by 4- or 8-year-olds that are false. Responses ranged from 0 to 90%; mean = 28%, median = 25%. Of the respondents, 15% estimated that 10% or fewer sexual abuse claims were false; 23% of the respondents' estimates fell between 10 and 20%; 26% fell between 25 and 30%; 11% fell between 35 and 49%; and 17% estimates were 50% or more. A 2 (child age) \times 2 (participant gender) \times 2 (sample) ANOVA revealed significant effects of child age and participant gender, $F_s(1, 260) \geq 8.42$, $p_s < .01$, $\eta^2 = .03$ and $.05$ for age and gender, respectively. Estimates of false reports were significantly higher for 4-year-olds, $M = 30\%$, than for 8-year-olds, $M = 24\%$. Also, males gave higher estimates of false reports, $M = 32\%$, than did females, $M = 23\%$. Both patterns are consistent with empirical research: younger children are more suggestible and hence may well be more prone to false reports of abuse, and men are more skeptical of children's sexual abuse claims than are women. Finally, decreases in individuals' overall belief accuracy were associated with somewhat larger estimates of false reports, $r(268) = -.12$, $p = .05$. Our respondents' answers are not as high as those reported by Morrison and Greene (1992), who found that jurors estimated that approximately half of all allegations are false. But they are higher than the rate of less than 10% suggested by research (e.g., Everson & Boat, 1989; Faller, Froning, & Lipovsky, 1991; Herbert, 1987; Jones & McGraw, 1987). Thus, whether or not participants believe a specific allegation of sexual abuse, they appear to believe that a substantial number of such allegations are false.

Knowledge of Memory and Suggestibility Studies

The final analyses concerned participants' knowledge of the results of four widely cited studies of children's memory, suggestibility, and false event reports. First, we compared participant estimates to the actual study results. We then examined how factors manipulated in the studies, sample, and participant gender related to participant estimates. Significant effects at $p \leq .01$ are reported.

Poole and Lindsay's (1995) Mr. Science study. When asked to estimate the number of children who erred and indicated that the man had put something yucky in their mouths, participants' responses ranged from 0 to 100%. On average, respondents overestimated the number of children who gave false reports in this study, with their errors being larger when estimating results concerning 3-year-olds in the first interview and estimating results concerning 8-year-olds in the second interview. Specifically, participants estimated that in the first interview 64% of the 3-year-olds said that Mr. Science put something yucky in their mouths, whereas, in the study, only 37% made this claim; $t(308) = 19.34$, $d = 1.10$, and that 40% of the 8-year-olds made this mistake, whereas in the study it was only 36%, $t(309) = 2.57$, $d = .15$. Participants estimated that in the second interview 54% of the 3-year-olds and 32% of the 8-year-olds made a mistake. In actuality, 54% of the 3-year-olds and 14% of the 8-year-olds did so. The difference between participants' estimates and the actual results in the second interview was only significant for the 8-year-olds, $t(306) = 11.55$, $d = .66$.

A 2 (sample: college student vs. jury pool member) \times 2 (gender) \times 2 (child age: 3- vs. 8-year-old) \times 2 (interview number: first vs. second) ANOVA was conducted with child age and interview number as within-subject factors. Participants' estimates of false reports were significantly higher for 3-year-olds than 8-year-olds, $F(1, 300) = 191.34$, $\eta^2 = .39$. This is consistent with other research and was in fact the case in this study—but only in the second interview. Participants' estimates of false reports were also significantly higher for the first than second interview, $F(1, 300) = 136.17$, $\eta^2 = .31$. This, too, was in fact the case—but only for the 8-year-olds. For older children, being told that some of the suggestions about Mr. Science were false decreased the likelihood of false reports. In general, our participants failed to recognize age differences in the benefits of interviewer feedback, which only helped the older children.

Ceci et al.'s (1994) mousetrap study. When participants estimated the percentage of children who falsely assented to fictitious events in the final interview in this study, estimates ranged from 0 to 100%. As was observed in the Poole and Lindsay study, participants again overestimated children's errors. For 3–4-year-olds, participants estimated that the children would falsely agree that they had experienced 52% of the false events; in the study it was only 36%; $t(307) = 12.15$, $d = .69$; and for 5–6-year-olds, they estimated an error rate of 40% when it was only 32%; $t(307) = 6.33$, $d = .36$.

A 2 (sample) \times 2 (gender) \times 2 (child age: 3–4 vs. 5–6 years) ANOVA, conducted with child age as a within-subject factor, revealed a significant main effect of child age, $F(1, 302) = 166.38$, $\eta^2 = .36$. Despite overestimating children's false report rates, participants recognized that the number of children led to make false reports about entirely fictitious events decreased with age.

Thompson et al.'s (1997) janitor study. In contrast to their judgments about the other studies, participants greatly underestimated the percentage of children who gave false reports in the janitor study—presumably because of the high percentage of children who were influenced by the suggestive statements in the study. In fact, only one child in the study (2% of the sample) answered all of the questions about the janitor's activities correctly, and none of the children answered all of the questions correctly in the follow-up interview with their parents. Participants estimated that 50% of the children were correct in the first interview and 51% were correct in the second interview, both of which were significantly higher than the actual percentages of 2% and 0%, $t_s(306 \text{ or } 307) \geq 32.40$, $d_s \geq 1.85$. Again, estimates ranged from 0 to 100%.

When a 2 (sample) \times 2 (gender) \times 2 (interview number: initial vs. follow-up) mixed factorial ANOVA was conducted, the sample \times interview number interaction was significant, as was the three-way interaction, $F_s(1, 303) > 9.03$, $p_s < .01$, $\eta^2_s = .03$. However, an evaluation of the confidence intervals failed to reveal any group differences across the three factors.

Saywitz et al.'s (1991) medical checkup study. After reading the description of this medical procedure study, participants estimated the percentage of girls who correctly and incorrectly reported vaginal touch in response to free-recall, anatomical doll, and leading questions. When estimates of girls' true reports of vaginal touch were examined, participants overestimated that 60% of the girls disclosed vaginal touch in free-recall and that 67% did so with the anatomical doll and props, compared with actual figures of 22 and 17%; $t_s(305) = 24.35$ and 34.13 , $d_s = 1.39$ and 2.01 , respectively. Yet, participants underestimated the percentage of girls who disclosed when asked directly about vaginal touch while the interviewer pointed to the doll's genitals, 77% rather than 86%, $t(305) = -7.03$, $d = .40$.

Participants' estimates of true disclosures were then subjected to a 2 (sample) \times 2 (gender) \times 3 (question type: free-recall, doll, doll plus leading) ANOVA with question type varied within subject. A significant main effect of question type revealed that participants' estimates of the percentage of girls who disclosed vaginal touch increased as the questions became more leading; $F(2, 604) = 97.45$, $\eta^2 = .24$.

When estimates of girls' false reports of vaginal touch were examined, consistent with the *Mr. Science* and *mousetrap* studies, participants overestimated children's inaccuracies. None of the girls falsely claimed to have been touched in free-recall or in response to free-recall questions asked in conjunction with the anatomical doll. When asked the leading question while the interviewer pointed to the doll's genitals, only one girl (3%) made a false report. Participants estimated that 20, 19, and 25% of the girls falsely claimed vaginal touch across the three question types, $t_s(305) = 15.60$, 16.23 , and 11.94 , $d_s = .88$, $.93$, $.94$, respectively.

Next, participants' false report estimates were entered into the 2 (sample) \times 2 (gender) \times 3 (question type) ANOVA. A significant main effect of gender revealed that males' estimates of false reports were significantly higher than were females' estimates, $M_s = 30\%$ versus 20% , respectively, $F(1, 301) = 21.94$, $\eta^2 = .07$. A significant main effect of question type revealed that, although participants' estimates of children's false reports were always higher than the actual results, participants were nonetheless correct in the direction of their estimates, specifically that

false reports increased as questions became more leading, $F(2, 602) = 22.93, \eta^2 = .07$.

Summary. Participants' knowledge of specific research findings was quite limited, and their estimates of children's accuracy in four key studies ran the gamut from 0 to 100%. For three of the studies (*Mr. Science*, *mousetrap*, and *medical checkup*), participants clearly recognized the potential for suggestibility, and their estimates of children's errors exceeded those actually observed. This pattern of results supports Lyon's (1999) belief that jurors are already aware that children are suggestible. Yet, in the fourth study (*janitor*), participants' estimates of child errors were substantially lower than those observed in the study—a finding that contradicts Lyon's (1999) claim.⁵ In this case, participants appeared not to recognize the powerful influence of stereotypic inductions on children's eyewitness accuracy (Thompson et al., 1997; see also Garven, Wood, Malpass, & Shaw, 1998; Leichtman & Ceci, 1994; Poole & Lindsay, 2002). It may be that, although individuals are knowledgeable and perhaps even overly skeptical about some aspects of children's suggestibility, they are less aware of adverse effects of somewhat subtle but highly influential interview manipulations. Thus, knowledge provided by an expert regarding situational influences on children's eyewitness reporting can be critical in assisting jurors as they evaluate the evidence in a case. It is also possible that participants in this study were simply guessing about study results. For three of the studies, their estimates of suggestibility averaged approximately 50%: across age and interview, their estimate of children's false memories in the *Mr. Science* study was 48%, in the *mousetrap* study was 46%, and in the *janitor* study was 50%. It was only for false reports of vaginal contact in the *medical checkup* study that the participants estimated a lower likelihood of false reports (21%). Information from an expert may be necessary to move jurors from guessing to knowing about the rates of false reports in children.

CONCLUSION

The goal of the present study was to examine jurors' and potential jurors' general beliefs concerning children's memory, suggestibility, and reactions to and disclosure of sexual abuse. The results are directly relevant to current debates regarding the necessity of expert witnesses in cases involving child victim/witnesses, particularly those alleged to have experienced sexual abuse (see Ceci & Friedman, 2000; Lyon, 1999). As elaborately put forth by Lyon (1999, 2002), if jurors already know about children's suggestibility, and about the factors that lead them to make errors, expert testimony regarding children's suggestibility is not necessary and may in fact lead to unwarranted levels of juror skepticism. Yet, Ceci and Friedman

⁵When participants were asked to estimate the proportion of children who erred in response to particular questions, they tended to overestimate the level of suggestibility. When they were asked to estimate the proportion of children who answered all questions correctly, they underestimated suggestibility. This difference in the phrasing of questions may explain some of the evident differences in participants' responses, although it is unlikely that the difference is completely responsible for the evident patterns of results.

(2000) commented that there is no reason to suspect that jurors know about children's suggestibility, particularly the influence of highly leading interview tactics on children's errors and false event reports. Nor is there evidence that jurors are overly skeptical regarding children's abuse claims.

Our study represents an important step in evaluating the need for expert witnesses in child sexual abuse cases and in cases where suggestibility concerns are raised. However, several limitations to the generalizability of our results need to be mentioned. First, an important strength of the study was its inclusion of individuals who just completed jury duty rather than only college students or randomly surveyed community members. Thus, our results are likely to generalize to many actual juries. However, the study was conducted in a single county in California. There is no reason to expect that jurors' beliefs would differ considerably across different counties and states, but research is needed to confirm this. Second, as mentioned, there were limitations in our ability to classify statements on the belief questionnaire as true or false. Our classifications were based on information collected from several sources, including empirical research, facts from legal cases, and prior expert opinion. Although we believe that scientists currently conducting research in this area would largely agree with our classifications, this is an empirical question worthy of direct investigation. Indeed, insofar as scientists are serving as experts in legal cases and describing children's eyewitness abilities, it is imperative to ascertain the extent to which they agree about the factors that inhibit or encourage children's allegations, disclosures, and reports of abuse. As more research is conducted, it is possible that some of our classifications will change. Thus, caution is advised for those who would attempt to generalize our findings to specific legal cases.

Despite the need for further research assessing both experts' perceptions and the generalizability of our findings, what can we tentatively conclude from our study? Do jurors already have requisite knowledge regarding children's memory, eyewitness abilities, and disclosure of abuse, or are jurors overly skeptical? Our results do not suggest a simple answer. Instead, the study revealed considerable variability in individuals' knowledge about children's eyewitness abilities and reactions to abuse and indicated that individuals possess both accurate and inaccurate beliefs. A critical and much needed role for experts is to reduce this variability and correct the misperceptions of a majority (or a large minority) of jurors. Indeed, Ceci and Friedman (2000) highlighted that, as long as a sizable minority of jurors hold incorrect beliefs, expert testimony is important in juror education. Our findings reveal that, even when a majority of individuals held correct beliefs, a large minority did not. Of importance, however, experts need to go beyond simply stating that children can be misled to make false claims of abuse and explain the conditions under which children are more (and less) likely to err. This point was made by Lyon (2002) who noted that experts' testimony may be useful in providing insight into specific research findings concerning the reliability of children's testimony that are not well understood by laypersons. Overall then, the knowledge provided by experts could reduce both unwarranted skepticism and naive trust in children's claims of sexual abuse.

ACKNOWLEDGMENTS

The study was funded by an interdisciplinary grant from the School of Social Ecology at the University of California, Irvine. We would like to thank Michael Kromer for his extensive assistance with data collection and coding. Gratitude is also expressed to Alan Slater, the Chief Executive Officer and Jury Commissioner of the Superior Court of Orange County, California and his staff, particularly Thu Nguyen, the Jury Manager. Without their help and support, this study would not have been possible.

REFERENCES

- Bornstein, B. H. (1999). The ecological validity of jury simulations: Is the jury still out? *Law and Human Behavior, 23*, 75–91.
- Bottoms, B. L., & Goodman, G. S. (1994). Perceptions of children's credibility in sexual assault cases. *Journal of Applied Social Psychology, 24*, 702–732.
- Brigham, J. C., Van Verst, M., & Bothwell, R. K. (1981). Accuracy of children's eyewitness identifications in a field setting. *Basic & Applied Social Psychology, 7*, 295–306.
- Brown, D., & Pipe, M.-E. (2003). Variations on a technique: Enhancing children's recall using narrative elaboration training. *Applied Cognitive Psychology, 17*, 377–399.
- Bruck, M., Ceci, S. J., Francoeur, E., & Renick, A. (1995). Anatomical dolls do not facilitate preschoolers' reports of a pediatric examination involving genital touching. *Journal of Experimental Psychology: Applied, 1*, 95–109.
- Bruck, M., Ceci, S. J., & Hembrooke, H. (2002). The nature of children's true and false narratives. *Developmental Review, 22*, 520–554.
- Bussey, K. (1992). Lying and truthfulness: Children's definitions, standards, and evaluative reactions. *Child Development, 63*, 129–137.
- Cacciola, E. R. (1986). The admissibility of expert testimony in intrafamilial child sexual abuse cases. *UCLA Law Review, 34*, 175.
- Carter, C. A., Bottoms, B. L., & Levine, M. (1996). Linguistic and socioemotional influences on the accuracy of children's reports. *Law and Human Behavior, 20*, 335–358.
- Ceci, S. J., & Bruck, M. (1993). Suggestibility of the child witness: A historical review and synthesis. *Psychological Bulletin, 113*, 403–439.
- Ceci, S. J., & Bruck, M. (1995). *Jeopardy in the courtroom: A scientific analysis of children's testimony*. Washington, DC: American Psychological Association.
- Ceci, S. J., & Bruck, M. (1998). Children's testimony: Applied and basic issues. In W. Damon, I. E. Sigel, & K. A. Renninger (Eds.), *Handbook of child psychology* (5th ed., pp. 713–773). New York: Wiley.
- Ceci, S. J., & Friedman, R. D. (2000). The suggestibility of children: Scientific research and legal implications. *Cornell Law Review, 86*, 33–108.
- Ceci, S. J., Huffman, M. L., Smith, E., & Loftus, E. F. (1994). Repeatedly thinking about a non-event: Source misattributions among preschoolers. *Consciousness & Cognition: An International Journal, 3*, 388–407.
- Ceci, S. J., Loftus, E. L., Leichtman, M. D., & Bruck, M. (1994). The possible role of source misattributions in the creation of false beliefs among preschoolers. *International Journal of Clinical & Experimental Hypnosis, 42*, 304–320.
- Cirelli, M. K. (1991). Expert testimony in child sexual abuse cases: Helpful or prejudicial. *Cooley Law Review, 8*, 425–449.
- Conte, J. R., & Schuerman, J. R. (1987). The effects of sexual abuse on children: A multidimensional view. *Journal of Interpersonal Violence, 2*, 380–390.
- Cooper, J., Bennett, E. A., & Sukel, H. L. (1996). Complex scientific testimony: How do jurors make decisions? *Law & Human Behavior, 20*, 379–394.
- Corder, B. F., & Whiteside, R. (1988). A survey of jurors' perception of issues related to child sexual abuse. *American Journal of Forensic Psychology, 6*, 37–43.
- Crowley, M. J., O'Callaghan, M. G., & Ball, P. J. (1994). The juridical impact of psychological expert testimony in a simulated child sexual abuse trial. *Law & Human Behavior, 18*, 89–105.

- Daubert v. Merrell-Dow Pharmaceuticals, 509 U.S. 579 (1993).
- Davis, S. L., & Bottoms, B. L. (2002). Effects of social support on children's eyewitness reports: A test of the underlying mechanism. *Law and Human Behavior, 26*, 185–215.
- DeJong, A. R., & Rose, M. (1991). Legal proof of child sexual abuse in the absence of physical evidence. *Pediatrics, 88*, 506–511.
- Elliot, D. M., & Briere, J. (1994). Forensic sexual abuse evaluations of older children: Disclosures and symptomatology. *Behavioral Sciences & the Law, 12*, 261–277.
- Everson, M. D., & Boat, B. W. (1989). False allegations of sexual abuse by children and adolescents. *Journal of the American Academy of Child & Adolescent Psychiatry, 28*, 230–235.
- Faller, K. C., Froning, M. L., & Lipovsky, J. (1991). The parent-child interview: Use in evaluating child allegations of sexual abuse by the parent. *American Journal of Orthopsychiatry, 61*, 552–557.
- Farrar, M., & Goodman, G. (1992). Developmental changes in event memory. *Child Development, 63*, 173–187.
- Fivush, R., Hudson, J., & Nelson, K. (1984). Children's long-term memory for a novel event: An exploratory study. *Merrill-Palmer Quarterly, 30*, 303–316.
- Fivush, R., & Schwarzmueller, A. (1995). Say it once again: Effects of repeated questions on children's event recall. *Journal of Traumatic Stress, 8*, 555–580.
- Fivush, R., & Schwarzmueller, A. (1998). Children remember childhood: Implications for childhood amnesia. *Applied Cognitive Psychology, 12*, 455–473.
- Flint, R. L. (1995). Child sexual abuse accommodation syndrome: Admissibility requirements. *American Journal of Criminal Law, 23*, 171–194.
- Frye v. United States, 293 F. 1013 (D.C. Circ. 1923).
- Garven, S., Wood, J. M., Malpass, R. S., & Shaw, J. S., III (1998). More than suggestion: The effect of interviewing techniques from the McMartin Preschool case. *Journal of Applied Psychology, 83*, 347–359.
- Gomes-Schwartz, B., Horowitz, J., & Sauzier, M. (1985). Severity of emotional distress among sexually abused preschool, school-age, and adolescent children. *Hospital and Community Psychiatry, 36*, 503–508.
- Goodman, G. S., Bottoms, B. L., Schwartz-Kenney, B., & Rudy, L. (1991). Children's memory for a stressful event: Improving children's reports. *Journal of Narrative and Life History, 1*, 69–99.
- Goodman, G. S., Quas, J. A., Batterman-Faunce, J. L., Riddlesberger, M., & Kuhn, G. (1997). Children's reactions to and memory for a stressful event: Influences of age, anatomical dolls, knowledge, and parental attachment. *Applied Developmental Science, 1*, 54–75.
- Goodman, G. S., Tobey, A., Batterman-Faunce, J., Orcutt, H., Thomas, S., Shapiro, C., & Sachsenmaier, T. (1998). Face-to-face confrontation: Effects of closed-circuit technology on children's eyewitness testimony and jurors' decisions. *Law and Human Behavior, 22*, 165–203.
- Gordon, S., & Jaudes, P. K. (1996). Sexual abuse evaluations in the emergency department: Is the history reliable? *Child Abuse & Neglect, 20*, 315–322.
- Gordon, B. L., Schroeder, & Adams, (1990a). Children's knowledge of sexuality: A comparison of sexually abused and nonabused children. *American Journal of Orthopsychiatry, 60*, 250–257.
- Gordon, B. L., Schroeder, & Adams, (1990b). Age and social-class differences in children's knowledge of sexuality. *Journal of Clinical Child Psychology, 19*, 33–43.
- Haugaard, J. J. (1993). Young children's classification of the corroboration of a false statement as the truth or a lie. *Law & Human Behavior, 17*, 645–659.
- Herbert, C. P. (1987). Expert medical assessment in determining the probability of alleged child sexual abuse. *Child Abuse and Neglect, 11*, 213–221.
- Hershkowitz, I., Orbach, Y., Lamb, M. E., Sternberg, K. J., Pipe, M.-E., & Horowitz, D. (2004, July). *Dynamics of forensic interviews with children who do not disclose abuse*. Paper presented at the Annual Convention of the American Psychological Association, Honolulu, HI.
- Howe, M. L., Courage, M. L., & Edison, S. C. (2003). When autobiographical memory begins. *Developmental Review, 23*, 471–494.
- Hudson, J., Fivush, R., & Kuebli, J. (1992). Scripts and episodes: The development of event memory. *Applied Cognitive Psychology, 6*, 483–505.
- Huffman, M. L., Crossman, A. M., & Ceci, S. J. (1997). Are false memories permanent? An investigation of the long-term effects of source misattributions. *Consciousness & Cognition: The International Journal, 6*, 482–490.
- Johnson, M. K., & Foley, M. A. (1984). Differentiating fact from fantasy: The reliability of children's memory. *Journal of Social Issues, 40*, 33–50.
- Jones, D. P., & McGraw, J. M. (1987). Reliable and fictitious accounts of sexual abuse to children. *Journal of Interpersonal Violence, 2*, 27–45.

- Kail, R. (1990). *The development of memory in children* (3rd ed.). New York: W. H. Freeman and Company.
- Kellogg, N. D., & Huston, R. L. (1995). Unwanted sexual experiences in adolescents: Patterns of disclosure. *Clinical Pediatrics*, *34*, 306–312.
- Kendall-Tackett, K. A., Williams, L. M., & Finkelhor, D. (1993). Impact of sexual abuse on children: A review and synthesis of recent empirical studies. *Psychological Bulletin*, *113*, 164–180.
- Kovera, M. B., & Borgida, E. (1997). Expert testimony in child sexual abuse trials: The admissibility of psychological science. *Applied Cognitive Psychology*, *11*, 105–129.
- Kovera, M. B., Gresham, A. W., Borgida, E., Grey, E., & Regan, P. C. (1997). Does expert psychological testimony inform or influence juror decision making? A social cognitive analysis. *Journal of Applied Psychology*, *82*, 178–191.
- Krauss, D. A., & Sales, B. D. (2001). The effects of clinical and scientific expert testimony on juror decision making in capital sentencing. *Psychology, Public Policy, & Law*, *7*, 267–310.
- Leichtman, M. D., & Ceci, S. J. (1994). The effects of stereotypes and suggestions on preschoolers' reports. *Developmental Psychology*, *31*, 568–578.
- Levy, R. J. (1989). Using 'scientific' testimony to prove child sexual abuse: The Dorsey & Witney professorship lecture. *Family Law Quarterly*, *23*, 383–409.
- Loftus, E. F. (2003). Make-believe memories. *American Psychologist*, *58*, 867–873.
- Loftus, E. F., & Pickrell, J. E. (1995). The formation of false memories. *Psychiatric Annals*, *25*, 720–725.
- Lucariello, J., & Nelson, K. (1985). Slot-filler categories as memory organizers for young children. *Developmental Psychology*, *21*, 272–282.
- Lyon, T. D. (1999). The new wave of suggestibility research: A critique. *Cornell Law Review*, *84*, 1004–1087.
- Lyon, T. D. (2002). Expert testimony on the suggestibility of children: Does it fit? In B. L. Bottoms, M. B. Kovera, & B. D. McCauliff (Eds.), *Children, social science, and the law* (pp. 378–411). Cambridge: Cambridge University Press.
- Lyon, T. D., & Saywitz, K. J. (1999). Young maltreated children's competence to take the oath. *Applied Developmental Science*, *3*, 16–27.
- McCauliff, B. D., & Kovera, M. B. (2005). Estimating the effects of misleading information on witness accuracy: Can expert testimony assist the trier of fact? Submitted manuscript.
- McCord, D. (1986). Expert psychological testimony about child complainants in sexual abuse prosecutions: A foray into the admissibility of novel psychological evidence. *Journal of Criminal Law & Criminology*, *77*, 1–68.
- Mannarino, T., & Cohen, J. (1986). A clinical-demographic study of sexually abused children. *Child Abuse and Neglect*, *10*, 17–23.
- Mian, M., Wehrspann, W. K., Kaljner-Diamond, H., LeBaron, D., & Winder, C. (1986). Review of 125 children 6 years of age and under who were sexually abused. *Child Abuse and Neglect*, *10*, 223–229.
- Monahan, J., & Walker, L. (1988). Social science research in law: A new paradigm. *American Psychologist*, *43*, 465–472.
- Morison, S., & Greene, E. (1992). Juror and expert knowledge of child sexual abuse. *Child Abuse & Neglect*, *16*, 595–613.
- Myers, J. E. B. (1993). Expert testimony regarding child sexual abuse. *Child Abuse & Neglect*, *17*, 175–185.
- Myers, J. E. B. (1997). *Evidence in child abuse and neglect cases* (3rd ed.). New York: Wiley.
- Myers, J. E. B., Bays, J., Becker, J., Berliner, L., Corwin, D. L., & Saywitz, K. J. (1989). Expert testimony in child sexual abuse litigation. *Nebraska Law Review*, *68*, 1–145.
- Ornstein, P. A., & Haden, C. (2002). The development of memory: Toward an understanding of children's testimony. In M. Eisen, J. A. Quas, & G. S. Goodman (Eds.), *Memory and suggestibility in the forensic interview* (pp. 29–62). Mahwah, NJ: Erlbaum.
- Peterson, C. (1999). Children's memory for medical emergencies: 2 years later. *Developmental Psychology*, *35*, 1493–1506.
- Poole, D. A., & Lamb, M. E. (1998). *Investigative interviews of children: A guide for helping professionals*. Washington, DC: American Psychological Association.
- Poole, D. A., & Lindsay, D. S. (1995). Interviewing preschoolers: Effects of nonsuggestive techniques, parental coaching, and leading questions on reports of nonexperienced events. *Journal of Experimental Child Psychology*, *60*, 129–154.
- Poole, D. A., & Lindsay, D. S. (2002). Children's suggestibility in the forensic context. In M. L. Eisen, J. A. Quas, & G. S. Goodman (Eds.), *Memory and suggestibility in the forensic interview* (pp. 355–382). Mahwah, NJ: Erlbaum.

- Quas, J. A., Goodman, G. S., Bidrose, B. L., Pipe, M.-E., Craw, S., & Ablin, D. (1999). Emotion and memory: Children's long-term remembering, forgetting, and suggestibility. *Journal of Experimental Child Psychology, 72*, 235–270.
- Quas, J. A., Bottoms, B. L., Haegerich, T., & Nysse-Carris, K. (2002). Effects of child and defendant gender on juror decision making in child sexual abuse trials. *Journal of Applied Social Psychology, 32*, 1993–2021.
- Quas, J. A., Goodman, G. S., Ghetti, S., & Redlich, A. (2000). Questioning the child witness: What can we conclude from the research thus far? *Trauma, Abuse, and Violence, 1*, 223–249.
- Quas, J. A., Malloy, L. M., Melinder, A., Goodman, G. S., Schaaf, J., & D'Mello, M. (2005). *Developmental differences in the effects of repeated interviews and leading instructions on young children's false-event reports*. Submitted manuscript.
- Quas, J. A., & Schaaf, J. M. (2002). Children's memories of experienced and nonexperienced events across repeated interviews. *Journal of Experimental Child Psychology, 83*, 304–338.
- Reyes, B. I. (2001). *A portrait of race and ethnicity in California: An assessment of social and economic well being*. San Francisco, CA: Public Policy Institute of California.
- Roe, R. J. (1985). Expert testimony in child sexual abuse cases. *University of Miami Law Review, 40*, 97–113.
- Rudy, L. A. (1996). *Custody and expert opinion as factors in sexual abuse cases*. Unpublished dissertation, State University of New York, Buffalo.
- 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. J., Goodman, G. S., Nicholas, E., & Moan, S. F. (1991). Children's memories of a physical examination involving genital touch: Implications for reports of child sexual abuse. *Journal of Consulting & Clinical Psychology, 59*, 682–691.
- Schneider, N. B. (1994). How the experts compare with judges and jurors in an awareness of child sex abuse factors and phenomena: Is there a role in court for the expert? Unpublished dissertation, Hahnemann University.
- Scullin, M. H., & Ceci, S. J. (2001). A suggestibility scale for children. *Personality and Individual Differences, 30*, 843–856.
- Sorensen, T., & Snow, B. (1986). How children tell: The process of disclosure in child sexual abuse. *Child Welfare, 70*, 3–15.
- State v. Sloan 912 S.W.2d 592 (Mo. Ct. App. 1995).
- Sternberg, K. J., Lamb, M. E., Esplin, P. W., & Baradaran, L. P. (1989). Using a scripted protocol in investigative interviews: A pilot study. *Applied Developmental Science, 3*, 70–76.
- Sternberg, K. J., Lamb, M. E., Esplin, P. W., Orbach, Y., & Herschkowitz, I. (2002). Using a structured interview protocol to improve the quality of investigative interviews. In M. Eisen, J. A. Quas, & G. S. Goodman (Eds.), *Memory and suggestibility in the forensic interview* (pp. 409–436). Mahwah, NJ: Erlbaum.
- Steward, M. S., & Steward, D. S. (1996). Interviewing young children about body touch and handling. *Monographs of the Society for Research in Child Development, Serial No. 248*.
- Summit, R. C. (1983). The child sexual abuse accommodation syndrome. *Child Abuse & Neglect, 7*, 177–193.
- Thompson, W. C., Clarke-Stewart, K. A., & Lepore, S. J. (1997). What did the janitor do? Suggestive interviewing and the accuracy of children's accounts. *Law & Human Behavior, 21*, 405–426.
- Underwager, R., & Wakefield, H. (1992). Poor psychology produces poor law. *Law & Human Behavior, 16*, 233–243.
- Usher, J. A., & Neisser, U. (1993). Childhood amnesia and the beginnings of memory for four early life events. *Journal of Experimental Psychology: General, 122*, 155–165.
- Warren, A. R., Woodall, C. E., Thomas, M., Nunno, M., Keeney, J. M., Larson, S. M., & Stadfeld, J. (1999). Assessing the effectiveness of a training program for interviewing child witnesses. *Applied Developmental Science, 3*, 128–135.
- West, T. A., & Bauer, P. J. (1999). Assumptions of infantile amnesia: Are there differences between early and later memories? *Memory, 7*, 257–278.