Cornell Law Review

Volume 84
Issue 4 *May 1999*

Article 3

New Wave in Children's Suggestibility Research: A Critique

Thomas D. Lyon

THE NEW WAVE IN CHILDREN'S SUGGESTIBILITY RESEARCH: A CRITIQUE

Thomas D. Lyon†

ABSTRACT

The new wave in children's suggestibility research consists of a prestigious group of researchers in developmental psychology who argue that children are highly vulnerable to suggestive interviewing techniques. Because of its scientific credentials, its moderate tone, and its impressive body of research, the new wave presents a serious challenge to those who have claimed that children are unlikely to allege sexual abuse falsely. Although we can learn much from the research, concerns over society's ability to detect abuse motivate three criticisms. First, the new-wave researchers assume that highly suggestive interviewing techniques are the norm in abuse investigations, despite little empirical evidence to support this claim. Second, the research neglects the characteristics of child sexual abuse that both make false allegations less likely and increase the need to guard against a failure to detect abuse when it actually has occurred. Third, the researchers' apparent value-free scientific treatment of the suggestibility issue obscures, rather than avoids, value judgments regarding the tradeoff between false allegations and false denials of sexual abuse.

† Associate Professor, University of Southern California Law School, Los Angeles, California. Email: tlyon@law.usc.edu. Thanks to Scott Altman, Mark Everson, Brian Holmgren, Michael Lamb, Martin Levine, Edward McCaffrey, Jeffrey Rachlinski, Elyn Saks, Karen Saywitz, Eric Talley, Michael Wald, Helen Westcott, and the participants at the University of Southern California Law School Faculty Workshop for their comments on drafts. Thanks to Verinder Shaw and to Hazel Lord and her staff at the University of Southern California Law Library for research assistance.

INTRODUCTION

For most of this century, psychologists and legal commentators have doubted the reliability of children's statements, particularly when those statements involve claims of sexual abuse. Although Sigmund Freud originally believed his adult patients' reports of childhood incest, his conviction that "surely such widespread perversions against children are not very probable"[1] led to his discovery of the oedipal complex, whereby young children generate incestuous fantasies about their opposite-sex parent.[2] Jean Piaget was not interested in sexual fantasies per se, but believed that the egocentric young child's thought was guided by imagination and unconstrained by reality: "[T]he child's mind is full of these 'ludistic' tendencies up to the age of 7-8, which means that before this age it is extremely difficult for him to distinguish between fabulation and truth."[3] In his classic treatise on evidence, John Henry Wigmore surveyed psychiatric evidence asserting that victims routinely fabricate allegations of sexual abuse and recommended that psychiatrists examine female complainants in sex-crime cases.[4]

---

1 Letter from Sigmund Freud to Wilhelm Fliess (Sept. 21, 1897), *in* THE COMPLETE LETTERS OF SIGMUND FREUD TO WILHELM FLIESS, 1887-1904, at 264 (Jeffrey Moussaieff Masson ed. & trans., 1985).

2 *See* SIGMUND FREUD, *The Transformations at Puberty*, *in* THREE ESSAYS ON THE THEORY OF SEXUALITY 93 (James Strachey ed. & trans., Hogarth Press Ltd. 1962) (1905) (referring to the "child's sexual impulses towards his parents, which are as a rule already differentiated owing to the attraction of the opposite sex—the son being drawn towards his mother and the daughter towards her father").

3 JEAN PIAGET, JUDGMENT AND REASONING IN THE CHILD 202 (Marjorie Warden trans., Littlefield, Adams 1972) (1928).

4 *See* 3A JOHN HENRY WIGMORE, EVIDENCE IN TRIALS AT COMMON LAW § 924a, at 737 (James H. Chadbourn ed., 1970). For more recent versions of the same view, see Hon. Charles F. Stafford, *The Child as a Witness*, 37 WASH. L. REV. 303, 309 (1962) (noting "the danger that a child will intermingle imagination with memory" in testimony); Note, United

Even critics of the giants of psychology often have assumed that children are wrong about sexual abuse. These critics merely move the source of the false allegation from the child to an influential adult. Some researchers claimed that the incest fantasies Freud uncovered were the product of his therapeutic method,[5] in which he had applied "the strongest compulsion" to overcome the "greatest reluctance" in patients to relate such fantasies.[6] Others believed the fantasies were attributable to ambivalent parental affection.[7] One could doubt Piaget on similar grounds. Some criticized Piaget's early methods of questioning children as too difficult and too suggestive.[8] Skeptics thus charged that Freud and Piaget had not "discovered" oedipal fantasies and childish egocentrism, but had invented phenomena that they then implanted in their subjects' heads. This view is consistent with 100 years of research and commentary on children's suggestibility. The research dates back at least as far as the turn of the century, when prominent psychologists such as Binet, Stern, and Varendonck warned courts about the dangers of children's testimony.[9]

In the 1970s, the women's movement persuaded researchers to take a new look at child sexual abuse.[10] Feminist writers reminded their readers that surveys revealed that one-fifth to one-third of adult

---

States v. Bear Runner: *The Need for Corroboration in Incest Cases*, 23 ST. LOUIS U. L.J. 747, 759-60 (1979) (asserting that children fantasize and lie about incest).

5 *See* RICHARD OFSHE & ETHAN WATTERS, MAKING MONSTERS: FALSE MEMORIES, PSYCHOTHERAPY, AND SEXUAL HYSTERIA 242 (1994) (arguing that Freud "bullied his patients in order that they might confirm his theories and interpretations").

6 3 SIGMUND FREUD, *The Aetiology of Hysteria*, *in* THE STANDARD EDITION OF THE COMPLETE PSYCHOLOGICAL WORKS OF SIGMUND FREUD 189, 204 (James Strachey et al. trans., Hogarth Press Ltd. 1962).

7 *See* NANCY CHODOROW, THE REPRODUCTION OF MOTHERING: PSYCHOANALYSIS AND THE SOCIOLOGY OF GENDER 160-61 (1978).

8 *See* MONIQUE LAURENDEAU & ADRIEN PINARD, CAUSAL THINKING IN THE CHILD 25 (1962) (discussing SUSAN ISAACS, INTELLECTUAL GROWTH IN YOUNG CHILDREN (1930) and JEAN M. DEUTSCHE, THE DEVELOPMENT OF CHIDREN'S CONCEPTS OF CAUSAL RELATIONS (1937)). For a more recent version of this criticism, see MARGARET DONALDSON, CHILDREN'S MINDS (1978).

9 *See* Stephen J. Ceci & Maggie Bruck, *Suggestibility of the Child Witness: A Historical Review and Synthesis*, 113 PSYCHOL. BULL. 403, 406-407 (1993) (discussing the work of Binet, Stern, and Varendonck); Jacqueline L. Cunningham, *Contributions to the History of Psychology: XLVI. The Pioneer Work of Alfred Binet on Children as Eyewitnesses*, 62 PSYCHOL. REP. 271, 271-76 (1988) (discussing the work of Binet and Stern); Gail S. Goodman, *Children's Testimony in Historical Perspective*, 40 J. SOC. ISSUES 9, 19-22 (1984) (discussing the work of Binet, Stern, and Varendonck). For a representative legal view, see M. RALPH BROWN, LEGAL PSYCHOLOGY 133 (1926) ("Create, if you will, an idea of what the child is to hear or see, and the child is very likely to hear or see what you desire.").

10 *See* JUDITH LEWIS HERMAN, FATHER-DAUGHTER INCEST 18 (1981). *See generally* SANDRA BUTLER, CONSPIRACY OF SILENCE: THE TRAUMA OF INCEST (1978) (explaining the dynamics of and society's responses to incestuous assault); BLAIR JUSTICE & RITA JUSTICE, THE BROKEN TABOO: SEX IN THE FAMILY (1979) (providing a primer on incest and suggestions for how to prevent it); KARIN C. MEISELMAN, INCEST (1978) (discussing the causes and effects of incest and providing recommendations for treatment).

women had some sort of sexual encounter with an adult male during childhood, and approximately fifteen percent had experienced abuse that involved physical contact.[11] A 1985 survey of a nationally representative sample of adult women and men found that twenty-seven percent of the women and sixteen percent of the men reported sexual abuse during childhood, and excluding noncontact abuse reduced the percentages only slightly.[12] These surveys proved that sexual abuse was more common than many people had imagined.

Spurred by growing awareness, legislators enacted legal reforms in the 1980s to facilitate the prosecution of child sexual abuse. Many states adopted special hearsay exceptions for children's allegations of abuse, and the courts broadly interpreted existing hearsay exceptions to admit a greater number of statements.[13] States also eliminated presumptions that children were testimonially incompetent.[14] As a result, prosecutors exhibited greater willingness to pursue child sexual abuse allegations.[15]

In this environment, the suggestibility of children re-emerged as an area of interest for researchers. Consistent with the new zeitgeist, researchers emphasized the accuracy of children's memories, particularly when recalling abuse. The leading figure in this movement was Gail Goodman, a developmental psychologist who was well-versed in laboratory research on memory development. Goodman challenged the traditional wisdom of suggestibility research by invoking the concept of ecological (or external) validity—the extent to which research applies to actual cases.[16] In the vast majority of studies examining children's suggestibility, researchers asked children questions about

---

11 *See* HERMAN, *supra* note 10, at 12-14 & tbl.1.1 (summarizing studies by Landis (1940); Kinsey (1953); Landis (1956); Gagnon (1965); and Finkelhor (1978)). The samples were comprised primarily of white, middle-class women. "In general, the poor, blacks and other minorities, rural people, and the mentally ill—those groups that are stereotypically suspected of deviant sexual activities—were conspicuous by their absence from these studies." *Id.* at 12.

12 *See* David Finkelhor et al., *Sexual Abuse in a National Survey of Adult Men and Women: Prevalence, Characteristics, and Risk Factors*, 14 CHILD ABUSE & NEGLECT 19, 20-21 & tbl.1 (1990).

13 *See* LUCY S. MCGOUGH, CHILD WITNESSES 15 (1994) (reporting that a number of states have adopted special hearsay exceptions for children's reports of abuse); CHRISTOPHER B. MUELLER & LAIRD C. KIRKPATRICK, MODERN EVIDENCE § 8.35, at 1218 & n.6, § 8.41, at 1264-66 & nn.15-16 (1995) (detailing the expanded use of medical diagnosis and excited utterance exceptions to admit children's reports of abuse).

14 *See* MCGOUGH, *supra* note 13, at 15.

15 *See* BARBARA E. SMITH & SHARON GORETSKY ELSTEIN, THE PROSECUTION OF CHILD SEXUAL AND PHYSICAL ABUSE CASES 26-27 & n.6 (1993) (stating that a telephone survey of 600 prosecutors nationwide revealed an increase in the number of sexual abuse cases prosecuted).

16 *See, e.g.*, Gail S. Goodman & Vicki S. Helgeson, *Child Sexual Assault: Children's Memory and the Law*, 40 U. MIAMI L. REV. 181, 185-207 (1985) (evaluating the efficacy of the legal system's current child questioning techniques in light of the psychological literature).

the peripheral details of trivial stimuli.[17] Demonstrations of suggestibility in these contexts are of dubious applicability to child abuse investigations because abuse investigators question children about the central details of their physical interactions with familiar adults.[18]

Goodman's research agenda entailed a test of children's suggestibility in contexts that she believed better approximated abuse investigations.[19] In a series of studies, she found that although there were age differences in suggestibility,[20] children were much less likely to assent falsely to questions related to physical or sexual abuse.[21] Her early research showed that children as young as four years of age were surprisingly resistant to suggestive questions implying abusive behavior. Young children rejected suggestions of abuse close to 100% of the time.[22]

Sensitive to issues of ecological validity, Goodman acknowledged that her research lacked many of the essential details of abuse investigations, including motivations to lie, suggestions to the child regarding the character of the accused, and repeated interviewing over a period of time.[23] Nevertheless, popularized summaries of her work often omitted any discussion of the limitations of her research.[24] Commentators often asserted that young children are no more sug-

---

17 *See id.* at 188.

18 *Cf. id.* at 188-89 ("It is likely to be more difficult to lead a child witness into making a false statement about a central piece of information.").

19 *See, e.g.*, Gail S. Goodman et al., *Child Sexual and Physical Abuse: Children's Testimony, in* CHILDREN'S EYEWITNESS MEMORY (Stephen J. Ceci et al. eds., 1987) [hereinafter Goodman et al., *Sexual and Physical Abuse*]; Gail S. Goodman et al., *The Child Victim's Testimony, in* NEW ISSUES FOR CHILD ADVOCATES (Ann M. Haralamhie ed., 1986) [hereinafter Goodman et al., *Victim's Testimony*]; Gail S. Goodman & Rebecca S. Reed, *Age Differences in Eyewitness Testimony*, 10 LAW & HUM. BEHAV. 317, 321 (1986).

20 Three-year-olds were particularly likely to assent to leading questions falsely. *See* Goodman et al., *Victim's Testimony, supra* note 19, at 167.

21 *See* Goodman et al., *Sexual and Physical Abuse, supra* note 19, at 17; Goodman et al., *Victim's Testimony, supra* note 19, at 167, 173; Goodman & Reed, *supra* note 19, at 324.

22 *See* Goodman et al., *Sexual and Physical Abuse, supra* note 19, at 17 (finding that three- to four-year-olds were more suggestible than five- to six-year-olds, but noting that both age groups were close to 100% accurate when responding to "Did the person kiss you?," "Did the person hit you?," and "Did the person put anything in your mouth?"); Goodman et al., *Victim's Testimony, supra* note 19, at 167, 170, 172 (finding that three-year-olds were more vulnerable than older children to leading questions, but noting that all cbildren correctly answered "no" to "Did the person hit you?" and "Did the person put anything in your mouth?"); Goodman & Reed, *supra* note 19, at 324 (finding three-year-olds were particularly vulnerable to leading questions, compared to six-year-olds and adults, but noting that subjects at each age were resistant to suggestions about "the central action").

23 *See* Gail S. Goodman et al., *Children's Concerns and Memory: Issues of Ecological Validity in the Stndy of Children's Eyewitness Memory, in* KNOWING AND REMEMBERING IN YOUNG CHILDREN 249, 280 (Robyn Fivusb & Judith A. Hudson eds., 1990).

24 *See, e.g.*, Anne C. Rourle, *Experts Fault McMartin Child Interview Methods*, L.A. TIMES, Jan. 25, 1990, at A1 ("Children, [Goodman and other] researchers insist, cannot tell a lie, at least not about something as painful and unfamiliar as sexual abuse.").

gestible than adults, a view that dovetailed nicely with an older claim that children do not—in fact cannot—lie about sexual abuse.[25]

The new-found faith in children's reliability proved to be short-lived. The first blow came from a spate of highly controversial allegations of abuse in daycare centers in the 1980s and early 1990s: Country Walk, McMartin, Fells Acres, Little Rascals, and Kelly Michaels.[26] Bizarre allegations of ritualistic abuse by preschool teachers became so incredible that they raised serious doubts regarding whether the abuse that the children reported was even possible, let alone provable beyond a reasonable doubt.[27] Critics focused attention on highly suggestive and sometimes overtly coercive interviewing by investigators,

---

25 Commentators who have argued that children are no more suggestible than adults include Laura Lane, Note, *The Effects of the Abolition of the Corroboration Requirement in Child Sexual Assault Cases*, 36 CATH. U. L. REV. 793, 806-07 (1987) ("[S]tudies show that children are no more suggestible than adults."); Andrea J. Weinerman, Note, *The Use and Misuse of Anatomically Correct Dolls in Child Sexual Abuse Evaluations: Uncovering Fact . . . Or Fantasy?*, 16 WOMEN'S RTS. L. REP. 347, 355 (1995) ("[R]ecent studies suggest that when adults and children are presented with misleading information, children are no more suggestible than adults."). For commentators arguing that children cannot or do not lie about sexual abuse, see Susan P. Mele, *Major Evidentiary Issues in Prosecutions of Family Abuse Cases*, 11 OHIO N.U. L. REV. 245, 267-68 (1984) (noting that the "small child . . . has yet to obtain a motive or desire to lie" and that "[t]he child's belief in adult omniscience, a God-like quality, impels the child to speak the truth" (emphasis and footnote omitted)); Roland Summit, *The Child Sexual Abuse Accommodation Syndrome*, 7 CHILD ABUSE & NEGLECT 177, 191 (1983) (referring to a "maxim among child sexual abuse intervention counselors and investigators that children never fabricate the kinds of explicit sexual manipulations they divulge in complaints or interrogations").

26 *See, e.g.*, Commonwealth v. Amirault, 677 N.E.2d 652, 674 (Mass. 1997) (denying the Fells Acres defendants' motion for a new trial); Commonwealth v. Amirault, 612 N.E.2d 631, 631 (Mass. 1993) (vacating trial judge's order to reduce the Fells Acres defendants' sentences and reinstating original sentences); Commonwealth v. LeFave, 556 N.E.2d 83, 93 (Mass. 1990) (affirming conviction of owner/teacher at Fells Acres for indecent assault and battery and rape of a child); Findings of Fact, Rulings of Law, and Order on Defendant's Motion for a New Trial, Commonwealth v. LeFave, No. 85-63 (Mass. Super. Ct. June 1998); State v. Michaels, 625 A.2d 489, 524 (N.J. Super. Ct. App. Div. 1993) (reversing and remanding nursery school teacher's conviction for sexual offenses involving children), *aff'd*, 642 A.2d 1372, 1385 (N.J. 1994); State v. Kelly, 456 S.E.2d 861, 869 (N.C. Ct. App. 1995) (ordering new trial after conviction of day care operator on 99 charges including first degree rape); PATRICIA CROWLEY, NOT MY CHILD: A MOTHER CONFRONTS HER CHILD'S SEXUAL ABUSE (1990) (supporting the prosecution of Michaels); PAUL EBERLE & SHIRLEY EBERLE, THE ABUSE OF INNOCENCE: THE MCMARTIN PRESCHOOL TRIAL (1993) (supporting McMartin defense); JAN HOLLINGSWORTH, UNSPEAKABLE ACTS (1986) (supporting Country Walk prosecution); LISA MANSHEL, NAP TIME (1990) (supporting the prosecution of Michaels); Debbie Nathan, *Revisiting Country Walk*, 5 ISSUES IN CHILD ABUSE ACCUSATIONS 1 (1993) (supporting Country Walk defense); *Frontline: Innocenee Lost* (PBS television broadcast, July 20-21, 1993) (supporting Little Rascals defense).

27 *See, e.g.*, EBERLE & EBERLE, *supra* note 26, at 171 (describing bizarre allegations in the *McMartin* case); Nathan, *supra* note 26, at 1 (describing "bizarre" allegations in the *Country Walk* case); *Frontline*, *supra* note 26 (describing bizarre allegations in the *Little Rascals* case).

therapists, and parents.[28] Doubts spread to children's allegations in less sensational cases.[29] If adults could mislead children to believe the unbelievable, then one reasonably could conclude that adults also could mislead children to believe in sexual abuse of a more mundane kind.

The scientific community delivered the second blow to a growing faith in children's abuse allegations. The daycare cases inspired a "new wave" of suggestibility research[30] that reinforces the conventional wisdom that children are highly suggestible. The new wave presents formidable qualifications. Stephen J. Ceci and his colleagues have performed the most visible research.[31] Ceci is a professor of psychology at Cornell University, is a well-respected researcher in memory development and intelligence, and is the author of over 150 articles, chapters, and books.[32] Maggie Bruck, a frequent collaborator of Ceci, is a professor of psychology at McGill University and has an extremely impressive research background in learning disabilities. Michelle Leichtman, a former student of Ceci and the first author of perhaps the best known of the new wave's studies, is now an assistant professor of psychology at Harvard University. Ceci and Bruck co-authored a comprehensive review of the past 100 years of research on children's suggestibility, which received the Society for the Psychological Study of Social Issues' award for the best article of the year on child abuse and received acclaim as "'an excellent example of how rigorous research can inform important social problems.'"[33] Ceci, Bruck, and their colleagues have published research on children's suggestibility in the most prestigious peer-reviewed psychology jour-

---

28 *See, e.g.*, EBERLE & EBERLE, *supra* note 26, at 200-02 (criticizing interviewing in the *McMartin* case); Nathan, *supra* note 26, at 2 (criticizing interviewing in the *Country Walk* case); *Frontline*, *supra* note 26 (criticizing interviewing in the *Little Rascals* case).

29 *See* SMITH & ELSTEIN, *supra* note 15, at 65-66 (noting that a survey of prosecutors handling child sexual abuse cases revealed that 39% perceived a "backlash by jurors due to publicity of cases such as the McMartin Preschool case").

30 Maggie Bruck et al., *Reliability and Credibility of Young Children's Reports*, 53 AM. PSYCHOL. 136, 137 (1998); *see also* Maggie Bruck & Stephen J. Ceci, *Issues in the Scientific Validation of Interviews with Young Children*, *in* INTERVIEWING YOUNG CHILDREN ABOUT BODY TOUCH AND HANDLING, at 204, 212 (Monographs of the Soc'y for Research in Child Dev. No. 248, 1996) ("The goal of the next wave . . . will need to focus on how empty the glass can get as conditions resemble those of actual forensic investigations: the focus will need to be on the assessment of the risks of these techniques as well on as their benefits.").

31 *See* Debra A. Poole & D. Stephen Lindsay, *Assessing the Accuracy of Young Children's Reports: Lessons from the Investigation of Child Sexual Abuse*, 7 APPLIED & PREVENTIVE PSYCHOL. 1, 3 (1998) (reviewing suggestibility research and noting that "[t]he most widely cited studies of children's suggestibility are those conducted by Stephen Ceci, Maggie Bruck, and their colleagues").

32 *See* Stephen J. Ceci, *Cognitive and Social Factors in Children's Testimony*, *in* PSYCHOLOGY IN LITIGATION AND LEGISLATION 11, 13-14 (Bruce D. Sales & Gary R. VandenBos eds., 1994).

33 Stephen J. Ceci et al., *Children's Allegations of Sexual Abuse: Forensic and Scientific Issues: A Reply to Commentators*, 1 PSYCHOL. PUB. POL'Y & L. 494, 499 (1995) (quoting the Memorial Award Committee of the American Psychological Association).

nals.[34] In February of 1998, Bruck, Ceci, and Helene Hembrooke published a review of children's suggestibility research in the American Psychologist,[35] a journal received by every member of the American Psychological Association.

Ceci and his colleagues also have written for legal audiences. In 1995, Ceci and Bruck co-authored *Jeopardy in the Courtroom: A Scientific Analysis of Children's Testimony,* published by the American Psychological Association.[36] The authors desigued the book principally for judges, attorneys, and others who work in the field of child protection.[37] As the title indicates, the new wave could not embrace more emphatically the scientific method as the means to discover truth. In both this book and their other writings, the authors emphasize the superiority of the scientific method over clinical experience or adversarial courtroom battles.[38] Ceci and Bruck's book marshals an impressive amount of research documenting the risk of false allegations arising from interviews with children about sexual abuse. Through a reanalysis of evidence once touted as proving children's resistance to suggestibility, and through carefully controlled research of their own, the authors provide a compelling picture of the potential dangers of suggestive abuse investigations.

Given its moderate tone and carefully stated conclusions, the work of the new wave likely will influence those in the courtroom who are interested in an impartial appraisal of a highly contentious field. Judges may look to the writings as background information that will shape their judgments regarding the admissibility and sufficiency of evidence in cases involving child sexual abuse. In 1993, after Kelly Michaels appealed her 1988 conviction to the New Jersey Supreme Court, Bruck and Ceci co-authored an amicus brief that reviewed the

---

34 *See, e.g.,* Maggie Bruck et al., *"I Hardly Cried When I Got My Shot!" Influencing Children's Reports About a Visit to Their Pediatrician,* 66 CHILD DEV. 193 (1995) [hereinafter Bruck et al., *I Hardly Cried*]; Maggie Bruck et al., *Anatomically Detailed Dolls Do Not Facilitate Preschoolers' Reports of a Pediatric Examination Involving Genital Touching,* 1 J. EXPERIMENTAL PSYCHOL.: APPLIED 95 (1995) [hereinafter Bruck et al., *Dolls*]; Stephen J. Ceci & Mary Lyn Crotteau Huffman, *How Suggestible Are Preschool Children? Cognitive and Social Factors,* 36 J. AM. ACAD. CHILD ADOLESCENT PSYCHIATRY 948 (1997); Michelle D. Leichtman & Stephen J. Ceci, *The Effects of Stereotypes and Suggestions on Preschoolers' Reports,* 31 DEVELOPMENTAL PSYCHOL. 568 (1995).

35 *See* Bruck et al., *supra* note 30.

36 STEPHEN J. CECI & MAGGIE BRUCK, JEOPARDY IN THE COURTROOM: A SCIENTIFIC ANALYSIS OF CHILDREN'S TESTIMONY (1995).

37 *See id.* at xii.

38 *See* CECI & BRUCK, *supra* note 36, at 302 (expressing a preference to base predictions on laboratory research rather than on "anecdotes, personal opinions, and ideological views about children's gullibility or innocence"); Ceci et al., *supra* note 33, at 496 (referring to criticisms of their position as "one-sided, emotional arguments that do not stand the test of scientific scrutiny"); *id.* at 501 (suggesting that a critic's "arguments fall flat because he uses tactics that although permissible in a courtroom, are inappropriate in the scientific arena").

research on children's suggestibility, much of it their own. Forty-three research psychologists co-signed the brief.[39] Only three of the researchers asked to sign refused to do so; Gail Goodman was one of the three.[40] Affirming the lower court's reversal of Michaels's conviction, the court adopted an unprecedented procedure whereby a court may prevent child witnesses from testifying due to suggestive pretrial questioning.[41] In 1998, at Cheryl Amirault LeFave's fourth appeal challenging her 1987 conviction in the Fells Acre molestation case, Maggie Bruck submitted an affidavit summarizing recent suggestibility research and testified about the research at a Massachusetts Superior Court hearing.[42] In June of 1998, the court held that the research Bruck described constituted "new evidence" proving that suggestive interviewing practices "forever tainted" the testimony of the child witnesses, necessitating a new trial at which the court would not allow the child witnesses to testify.[43]

As a result of rulings such as these, attorneys surely will refer to the new wave research in arguing motions regarding evidence, in questioning child interviewers, and in both selecting and questioning expert witnesses regarding the suggestibility of children. Ceci and Bruck have themselves served as expert witnesses in a few cases,[44] and more seasoned expert witnesses have referenced their research.[45] To the extent that the new wave reaches a larger audience—in part because of intense media coverage[46]—it likely will influence the attitudes of legislators who consider procedural modifications designed

---

39 *See* Ceci et al., *supra* note 33, at 498. The signatories include several of the nation's most well-respected researchers in psychology, such as Ulric Neisser and Paul Ekman. *See* Maggie Bruck & Stephen J. Ceci, *Amicus Brief for the Case of* State of New Jersey v. Michaels *Presented by Committee of Concerned Social Scientists*, 1 PSYCHOL. PUB. POL'Y & L. 272, 312-314 (1995).

40 *See* Interview with Gail S. Goodman, Professor of Psychology, University of California at Davis, in San Diego, Cal. (January 1997).

41 If a criminal defendant demonstrates a substantial likelihood that a child witness's testimony was the product of pretrial suggestion, the child cannot testify unless the state provides clear and convincing evidence that the testimony is reliable. *See* State v. Michaels, 642 A.2d 1372 (N.J. 1994).

42 *See* Findings of Fact, Rulings of Law, and Order on Defendant's Motion for a New Trial at 65, Commonwealth v. LeFave, No. 85-63 (Mass. Super. Ct. June 1998).

43 *See id.*

44 *See* Martinez-Macias v. Collins, 810 F. Supp. 782, 813 n.68 (W.D. Tex. 1991) (summarizing testimony of Stephen J. Ceci); Maggie Bruck, *The Trials and Tribulations of a Novice Expert Witness*, *in* EXPERT WITNESSES IN CHILD ABUSE CASES 85 (Stephen J. Ceci & Helene Hembrooke eds., 1998) (summarizing her testimony in the *Little Rascals* case and the *Martensville* trials); Deposition of Stephen J. Ceci, State v. Fijnje, No. 89-43952 (Fla. Cir. Ct. Sept. 24, 1990).

45 *See* United States v. Rouse, 100 F.3d 560, 582 (8th Cir. 1996) (Loker, J., dissenting) (noting that expert witness referred to work by Ceci and Bruck).

46 *See, e.g.*, *Morning Edition: Children's Memories* (NPR radio broadcast, June 26, 1997), *available in* 1997 WL 12821747; *20/20: From the Mouths of Babes* (ABC television broadcast, Oct. 22, 1993).

to facilitate or restrict the acceptance of child testimony, and perhaps even will influence lay people called to sit as jurors in sexual abuse trials.

Because of the new wave's potential influence, and because few psychologists and legal commentators have questioned its claims,[47] it deserves critical examination. As in any area in which science is called into service to set policy, one can challenge the impartiality of the new wave. Arguments labeled as "scientific" often fail to avoid subjective judgment, and more dangerously, tend to obscure subjectivity when it occurs.

This Article seeks to explore and challenge the often unstated factual assumptions and value judgments made by the new wave of suggestibility research. The new wave bases its research and its arguments on unproven factual assumptions about abuse investigations and allegations. The new wave presumes highly suggestive interviewing techniques are commonplace, based on an unrepresentative review of abuse investigations. It emphasizes cases in which multiple numbers of preschool children accuse day care providers of bizarre acts, presenting a distorted picture of the suggestibility problems in the typical case, in which interviews likely are less coercive and children are less vulnerable to suggestion.

With respect to the value judgments, I focus on the fact that inherent tradeoffs exist between two types of errors—false positives and false negatives—and note that the new wave emphasizes the risk of

---

47 *See* McGOUGH, *supra* note 13, at 65-76; Angela R. Dunn, *Questioning the Reliability of Children's Testimony: An Examination of the Problematic Elements*, 19 LAW & PSYCHOL. REV. 203 (1995); Sheila Taub, *The Legal Treatment of Recovered Memories of Child Sexual Abuse*, 17 J. LEGAL MED. 183, 190-91 (1996); Nancy E. Walker & Matthew Nguyen, *Interviewing the Child Witness: The Do's and the Don't's, the How's and the Why's*, 29 CREIGHTON L. REV. 1587, 1590, 1600-02 (1996); Dana D. Anderson, Note, *Assessing the Reliability of Child Testimony in Sexual Abuse Cases*, 69 S. CAL. L. REV. 2117, 2137-39 (1996); Robert G. Marks, Note, *Should We Believe the People Who Believe the Children?: The Need for a New Sexual Abuse Tender Years Hearsay Exception Statute*, 32 HARV. J. ON LEGIS. 207, 222 n.75 (1995). For other recent legal commentary warning of children's suggestibility, see John R. Christiansen, *The Testimony of Child Witnesses: Fact, Fantasy, and the Influence of Pretrial Interviews*, 62 WASH. L. REV. 705, 708-13 (1987); Thomas L. Feher, *The Alleged Molestation Victim, the Rules of Evidence, and the Constitution: Should Children Really Be Seen and Not Heard?*, 14 AM. J. CRIM. L. 227, 231-33 (1987); Julie A. Dale, Comment, *Ensuring Reliable Testimony from Child Witnesses in Sexual Abuse Cases: Applying Social Science Evidence to a New Fact-Finding Method*, 57 ALB. L. REV. 187, 195-99 (1993); Diana Younts, Note, *Evaluating and Admitting Expert Opinion Testimony in Child Sexual Abuse Prosecutions*, 41 Duke L.J. 691, 723-29 (1991). For commentary critical of the new wave of research, see Thomas D. Lyon, *False Allegations and False Denials in Child Sexual Abuse*, 1 PSYCHOL. PUB. POL'Y & L. 429, 430-36 (1995) (criticizing Ceci and colleagues' work); Lisa Manshel, *The Child Witness and the Presumption of Authenticity After* State v. Michaels, 26 SETON HALL L. REV. 685, 691-93 (1996) (same); John E.B. Myers et al., *Psychological Research on Children as Witnesses: Practical Implications for Forensic Interviews and Courtroom Testimony*, 28 PAC. L.J. 3, 18-20 (1996) (criticizing one of Ceci and colleagues' studies); Helen Westcott, *Jeopardy in the Courtroom: A Scientific Analysis of Children's Testimony*, 89 BRIT. J. PSYCHOL. (forthcoming 1999) (book review).

false positives in the design and interpretation of their results. Given my own value preferences—an acute awareness of true cases of abuse and the difficulty abused children have in revealing abuse—I emphasize the effects of children's fear, loyalty, and embarrassment. These factors not only increase the likelihood of false negatives, but also reduce the likelihood of false positives in the cases that one most often sees in court—allegations of abuse against people close to the child.

Part I discusses the new wave's critique of Goodman's and her colleagues' research claims that children are surprisingly invulnerable to suggestion. The critique reveals the factual presuppositions and value preferences of the new wave's research program. Part II outlines the leading studies the new wave has conducted in anticipation of the critique that follows.

Part III begins the critique by discussing the real world of sexual abuse investigations. This Part emphasizes the importance of examining real-world interviewing to determine the extent to which one can apply the new wave's research to actual cases. The new wave emphasizes atypical cases—those in which investigators question large numbers of preschool children about the actions of day care providers. Investigators are less likely to use suggestive techniques in the typical abuse case—one involving a single victim and an alleged offender who is close to the child and her family. The new wave utilizes suggestive methods that have not been documented as prevalent among real-world interviews. Although research examining investigative interviews finds that large numbers of "leading" questions are asked, analysis of how "leading" is defined reveals that the new wave's research employs far more suggestive questions. Finally, the new wave's research has not adequately explored challenges child witnesses face in court.

Part IV discusses the real world of child sexual abuse and outlines reasons why children might deny abuse, including fear, loyalty, and embarrassment. Moreover, young children's recall is deficient, and more direct questions are necessary to tap recognition. These factors both support the limited use of "leading" questions and decrease the likelihood that false allegations will occur when investigators ask such questions.

Part V examines the value judgments underlying the new wave's research. This Part discusses the way in which the new wave positions itself as objective and scientific, thus appearing more credible than the veteran defense experts in child sexual abuse. I argue, however, that the objectivity is more apparent than real and discuss the role of value judgments in recommending investigative methods and in focusing on the possibility (rather than the relative probability) of false allegations of abuse. I conclude that one must recognize the empiri-

cal limitations and value judgments of the new wave in order to evaluate fairly children's sexual abuse allegations.

I
THE NEW WAVE'S CRITIQUE OF GOODMAN

> *Certainly, based on what we know, we can "rig" experiments to support our pet theories about children, but this approach does little to further our understanding of actual child witnesses. What it suggests instead is that the biases of researchers rather than the credibility of children should be investigated.*
>
> —Stephen J. Ceci et al.[48]

Gail Goodman became the researcher-heroine of the child protection movement in the 1980s because her research supported claims that false allegations of abuse rarely, if ever, occur.[49] To lay the groundwork for the new wave of research, Ceci and Bruck critiqued Goodman's work in three ways. First, the new wave faulted Goodman's research on the same grounds that she criticized research before her: a lack of ecological validity.[50] The new wave emphasized that Goodman's work involved interviewers who typically asked leading questions only once, in a single interview, without strong motivations on the part of either the interviewer or the child to report nonevents falsely.[51] Second, the new wave faulted Goodman for selectively interpreting findings to support a favored position. For example, Ceci and Bruck noted that Goodman focused on particular questions that do not show age differences, rather than discussing suggestibility in general, which tends to decrease with age.[52] Third, the new wave criticized Goodman's claim that false affirmation rates to abuse-related questions are surprisingly low among young children: "Ironically, studies by Goodman and her colleagues provide some of

---

48 Stephen J. Ceci et al., *The Suggestibility of Children's Recollections, in* CHILD ABUSE, CHILD DEVELOPMENT, AND SOCIAL POLICY 117, 133 (Dante Cicchetti & Sheree L. Toth eds., 1993).

49 Indeed, Ceci and Bruck have suggested that "[p]erhaps no researcher has done more to redress the historical imbalance in favor of child witnesses than Gail Goodman. After almost a century of research criticizing and belittling the accuracy and suggestibility of child witnesses, Goodman has presented a far more optimistic picture of children's abilities." Ceci & Bruck, *supra* note 9, at 410.

50 *See* CECI & BRUCK, *supra* note 36, at 73-74.

51 *See id.*

52 *See* Ceci & Bruck, *supra* note 9, at 432 ("For example, although Goodman and her colleagues chose to focus on segments of their data that did not contain age differences (e.g., abuse-related suggestions, stress induction), . . . they almost always found age differences in overall suggestibility, with the youngest preschoolers being disproportionately more suggestible than older children."); *see also id.* at 410 n.2 (discussing Leslie Rudy and Gail S. Goodman, *Effects of Participation on Children's Reports: Implications for Children's Testimony*, 27 DEVELOPMENTAL PSYCHOL. 527 (1991), and indicating that "their conclusions concerning the effects of participation seem overgenerous, given the actual pattern of results").

the most compelling evidence that young children do in fact make false claims about actions, central events, and, even events that could be construed as being sexually abusive."[53]

New wave researchers emphasize that their criticism of Goodman's work is not a personal attack. They explain that "[t]here is nothing emotional or ad hominem" in their critique, but rather, "it is the very essence of what scientists consider to be their responsibility: to refute or reanalyze the findings of others."[54] The scientist's goal is to conduct and to interpret research without imposing her own values.[55]

The new wave is equally cognizant, however, that value-free interpretation of research is an ideal rather than a reality. Ceci warned that "some of the better known figures in this area of research have exhibited a partisanship that prods them to discuss their findings without making clear the limits and alternate interpretations."[56] Moreover, Ceci co-authored several studies demonstrating that "seemingly objective scientific criteria" may be invoked to criticize proposals "whose real offense might be their social and political distastefulness."[57] Understanding a scientist's value preferences is therefore useful when assessing his scientific critiques.

Unfortunately, explicit acknowledgment of psychologists' values potentially undermines the respect psychology receives (or hopes to receive) as an objective science. Maggie Bruck wrote that testifying as an expert witness in court taught her

> what a mistake it is to confuse researchers with research. Although it is easy to do because one comes to represent a "researcher" with a specific point of view, this is a mistake in terms of the profession's applied image. You should try your best in the courtroom not to talk about researchers; rather, you should only talk about studies.[58]

---

53 *Id.* at 433.

54 Ceci et al., *supra* note 33, at 502.

55 *See id.* at 504 n.6 (noting that although they "readily accept the rarity of value-free observations" in science, they believe that "[r]esponsible scientists do not abandon their methodology, when it comes to the interpretation stage, but apply it with equal force and relevance").

56 Ceci, *supra* note 32, at 45.

57 Stephen J. Ceci et al., *Human Subjects Review, Personal Values, and the Regulation of Social Science Research*, 40 AM. PSYCHOL. 994, 1001 (1985); *see also* Douglas P. Peters & Stephen J. Ceci, *Peer-Review Practices of Psychological Journals: The Fate of Published Articles, Submitted Again*, 5 BEHAV. & BRAIN SCI. 187, 189-91 (1982) (discussing their study in which they submitted for publication 12 previously-published research articles authored by researchers from prestigious psychology departments, substituting names of unknown researchers from less-prestigious schools). The study found that 89% of the reviewers recommended rejection. *See id.* at 189. "[T]he manuscripts were rejected primarily for reasons of methodology and statistical treatment, not because reviewers judged that the work was not new." *Id.* at 191.

58 Bruck, *supra* note 44, at 95.

What is bad for experimental psychology's "applied image," however, is good for understanding its limits. Indeed, examination of Bruck's testimony in the cases she discusses clearly illustrates the differences in factual assumptions and value preferences among suggestibility researchers.

Bruck criticized at least three different studies in which Goodman collaborated.[59] This Article discusses two here. First, consider a study by Saywitz, Goodman, and their colleagues.[60] The researchers examined seventy-two five- and seven-year-old girls' memories of a pediatric examination.[61] For half of the girls, the examination included genital touch (exterior vaginal and anal examination), and for the other half, the examiner substituted an examination for scoliosis.[62] The study found that, with respect to girls' subsequent recall of genital touch, both false positive and false negative rates varied depending on the manner in which the examiner questioned the girls. When asked free-recall questions about the event either one week or one month afterwards, none of the girls in the scoliosis condition falsely claimed to have been touched in the genital area.[63] Of the girls in the genital touch condition, twenty-two percent (8/36) correctly mentioned vaginal touch, and eleven percent (4/36) correctly mentioned anal touch.[64] Free recall thus elicited no false positives but a substantial number of false negatives. When asked a direct question about genital touch with the aid of an anatomically correct doll (e.g., "Did that doctor touch you there?" while "pointing to the doll's vagina"),[65] 2.86% (1/35) of the girls in the scoliosis condition falsely claimed vaginal touch,[66] and 5.56% (2/36) falsely claimed anal touch.[67] Examiners asked the three girls who falsely claimed genital touch follow-up questions, and "two were unable to provide any detail. However, one

59 *See* Testimony of Maggie Bruck at 15,429-30, State v. Kelly, No. 933SC676 (N.C. Super. Ct. 1992) (describing testimony of Bruck who criticized Gail Goodman et al., Children's Testimony Nearly Four Years After an Event (1989) (unpublished manuscript), a paper presented at the meeting of the Eastern Psychology Association, Boston, MA, which is summarized in Gail S. Goodman & Alison Clarke-Stewart, *Suggestibility in Children's Testimony: Implications for Sexual Abuse Investigations*, *in* THE SUGGESTIBILITY OF CHILDREN'S RECOLLECTIONS 92 (John Doris ed., 1991)); *id.* at 15,462 (criticizing Karen J. Saywitz et al., *Children's Memories of a Physical Examination Involving Genital Touch: Implications for Reports of Child Sexual Abuse*, 59 J. CONSULTING & CLINICAL PSYCHOL. 682 (1991)); Testimony of Maggie Bruck at 5251, The Queen v. Sterling, Q.B.J. No. 74 (Sask.) (1994) (criticizing Rudy & Goodman, *supra* note 52, at 527).

60 *See* Saywitz et al., *supra* note 59.

61 *See id.* at 683.

62 *See id.* at 684.

63 *See id.* at 686.

64 *See id.*

65 *Id.* at 684 (internal quotation marks omitted).

66 *See id.* at 687 ("The 2.86% rate is based on 35 children because one parent crossed out the vaginal touch question.").

67 *See id.*

child in the nongenital condition who said yes to the anal touch question described in further questioning that 'it tickled' and 'the doctor used a long stick.'"[68] Of the girls in the genital condition, eighty-six percent (31/36) acknowledged vaginal touch when directly asked, and sixty-nine percent (25/36) acknowledged anal touch.[69] In comparison to free recall, a direct question elicited some false positives, but reduced the number of false negatives.

In her testimony, Bruck criticized the study on two grounds. First, she contended that "[i]t's a meaningless study. Those kids were questioned in totally unrealistic ways in terms of what goes on in sexual abuse cases."[70] Second, she disagreed with the authors' assessment of the significance of the false affirmations of touching:

> [T]here were three children in her study who made incredible claims about being touched. They kind of buried that under the data. I think that those cases are really important because here these children were only interviewed once and three children claimed that they had been touched in the genitals, one child claims that the doctor had shoved something up her hiney. I find that highly significant. She doesn't; I do.[71]

In their paper, Saywitz, Goodman, and their colleagues acknowledged that they did not repeatedly interview the children and that children in the nongenital condition "had no motive to distort their reports," thus potentially underestimating the suggestibility of children in forensic interviews.[72] However, they claimed "greater ecological validity" for their study than previous research and believed that their data would be useful to clinicians and legal professionals weighing the costs and benefits of different interviewing strategies.[73] The researchers also acknowledged that a cost-benefit analysis would require one to consider the children who made false allegations of vaginal and anal touch. They argued, however, that "although there is a risk of increased error with doll-aided direct questions, there is an *even greater risk* that not asking about vaginal and anal touch leaves the majority of such touch unreported."[74]

---

68 *Id.*

69 *See id.* at 686-87.

70 Testimony of Maggie Bruck at 15,458-59, State v. Kelly, No. 933SC676 (N.C. Super. Ct. 1992); *see also* Testimony of Maggie Bruck at 5321, The Queen v. Sterling, Q.B.J. No. 74 (Sask.) (1994) (explaining that Bruck failed to discuss the Saywitz study in her direct examination because of her belief that the study is not relevant to cases in which children are given repeated interviews with repeated suggestions).

71 Testimony of Maggie Bruck at 15,462, *Kelly*, No. 933SC676.

72 Saywitz et al., *supra* note 59, at 691.

73 *Id.* at 690.

74 *Id.* (emphasis added). Technically, the authors are correct in stating that the risk of unreported touch is greater than the risk of falsely reported touch, because the percentage of touched girls who failed to disclose is greater than the percentage of untouched

Second, consider a study by Goodman and her colleagues which Bruck also discussed in her testimony in the *Kelly* trial.[75] Fifteen seven- and ten-year-old children were interviewed four years after a five-minute interaction with an unfamiliar male adult. The interviewers created an "atmosphere of accusation" by interrogating the children with suggestive questions and comments (e.g., "You'll feel better once you've told" and "Are you afraid to tell?"),[76] and by asking abuse-related questions (e.g., "Did he do anything that made you feel uncomfortable?" and "He gave you a hug and kissed you, didn't he?").[77] According to Goodman:

> In free recall, few children evidenced memory of the original experience. They made a variety of errors in attempting to recall the event and answer questions. The children did not, however, provide false reports of abuse. All of the children knew their clothes had remained on, they had not been touched in a bad way, they had not been spanked, and they had not been instructed to keep a secret. Some of the children's errors, however, might lead to suspicion of abuse. For example, one child falsely affirmed that she had been given a bath, five children agreed to having been both hugged and kissed, and two children said "yes" when asked if their picture had been taken in the bathtub. Nevertheless, the children were more resistant to abuse-related than to nonabuse-related suggestions.[78]

Bruck described the study in her direct examination at the *Kelly* trial:

> And what was really surprising in this study was that almost a third of the children, in fact, claimed that the following things had happened to them: That they had been hugged or kissed when, in fact, [they] had never been hugged or kissed; that they had been taken into a bathroom when that never had happened; that they, in fact, had been taken into a bathtub when that never had happened; and one of the children actually claimed to have been given a bath.
>
> Now, this study is quite powerful because it shows how with just very slight manipulations and in one interview situation you can get children who had no memories for an event [to] start to say that certain things happened.

---

girls who falsely claimed they were touched. Yet calling nondisclosure the "greater risk" implicitly assumes that both kinds of risk are weighed equally. Moreover, the calculations assume that 50% of children questioned have in fact been touched. If the percentage is much lower than 50%, even a very low error rate may mean that direct questions lead to reports of genital touch that are more likely false than true. These issues are discussed at greater length *infra* text accompanying notes 410-12.

75 *See* Goodman et al., *supra* note 59.

76 Testimony of Maggie Bruck at 15,310, *Kelly*, No. 933SC676 (internal quotation marks omitted).

77 Goodman & Clarke-Stewart, *supra* note 59, at 97 (internal quotation marks omitted).

78 *Id.* at 97-98.

> And the motivations that are provided in this study are really similar [to] ones that we see in lots of other kinds of cases, such as, the way to keep safe is to tell the bad secret, the way to get rid of bad feelings in my tummy is to tell my mommy, the more I tell the more they won't get me.[79]

The prosecutor challenged Bruck's interpretation of the study with a quotation from the original report, which argued that "the children were surprisingly accurate in knowing that their clothes had remained on, that they had not been spanked, that they had not been touched in a place where they didn't like it, and that they had not been instructed to keep a secret."[80] Bruck responded by saying

> That is straight out of their conclusions .... It does not match very well with their data. This study is a prime example of—a very important example of how researchers can collect certain kinds of data and look at them and report them in ways that are not there.
>
> Gail Goodman is a renown[ed] researcher who has an incredible bias. She collects very important data that consistently shows that young children are suggestible, and yet in terms of who's known in the scientific community, people say Gail Goodman, she's the one who does all the studies to show that children aren't suggestible.[81]

Reiterating the results she discussed on direct, Bruck continued: "Gail Goodman feels that that's not really very significant. I happen to feel, most scientists happen to feel, most people in forensic psychology or anybody who is involved in a case—court case happen to feel this is really significant."[82]

Bruck's testimony exemplifies each of the predominant criticisms of Goodman's work found in Ceci and Bruck's published work. She argued that real-world investigative interviewing is unlike that the Saywitz study used, making its findings meaningless. She also asserted that both the Saywitz and the Goodman studies chose to focus only on aspects of their data, thus concealing the actual findings. Finally, she contended that the research suggests a great danger that false positives will occur, rather than reassuring us about the credibility of children's claims.

The criticisms illustrate the role that psychologists' assumptions about reality and personal standards of proof play in affecting their evaluation of research. In dismissing the Saywitz study as meaningless, the new wave makes a factual claim about the nature of investigative interviewing. Until researchers have adequately canvassed actual in-

---

79 Testimony of Maggie Bruck at 15,311, *Kelly*, No. 933SC676.

80 *Id.* at 15,429.

81 *Id.* at 15,429-30.

82 *Id.* at 15,431.

terviews for suggestiveness, opinions about real-world interviewing entail subjective judgments based on limited personal experience. In criticizing Goodman's selective focus on particular results, such as low rates of false positives on questions that directly implicate abuse, the new wave challenges Goodman's assumption that children are much less likely to affirm abusive experiences falsely than other types of experiences. This challenge is a factual claim. Even if supported by research, however, the claim leaves room for disagreement among researchers regarding whether research that does not examine allegations of abuse is nevertheless useful in assessing the suggestibility of children in abuse cases. Finally, the new wave argues that the magnitude of errors in Goodman's research is shockingly high, rather than reassuringly low, thereby expressing a value judgment regarding how many false allegations we are willing to accept in order to identify true cases of abuse. This Article takes up each of these points in the review of the new wave's research that follows.

II
STUDIES OF THE NEW WAVE

When legal commentators discuss the work of the new wave, they tend to accept its conclusions at face value[83] in much the same way that legal commentators eagerly touted earlier work on suggestibility purporting to prove that children are not suggestible.[84] This is a mistake. In her testimony in the *Kelly* case, Bruck warned the prosecutor that "[y]ou can't look at the conclusions. You have to look at the data."[85] Conclusions are a product of the results *and* the subjective impressions of the researcher. To go one step further, one must remember that researchers' assumptions about the world and their value preferences also affect the design of the research. Although the rigors of the scientific method often thwart even the cleverest researchers and the most brilliant hypotheses, one should recall Ceci's admonishment that researchers "can 'rig' experiments to support [their] pet theories about children."[86]

The new wave has produced dozens of studies in the past few years, including four studies that are particularly noteworthy.

---

83 *See supra* note 47 (citing law review articles favorably mentioning new wave research).

84 *See supra* note 25 (citing law reviews arguing children are no more suggestible than adults).

85 Testimony of Maggie Bruck at 15,433, *Kelly*, No. 933SC676.

86 Ceci et al., *supra* note 48, at 133.

A. Leichtman and Ceci's Sam Stone Study

In Leichtman and Ceci's Sam Stone Study, research assistants visited preschool children once a week for four weeks and told them about twelve incidents involving a clumsy fellow named Sam Stone.[87] Subsequently, Sam Stone visited the classroom while the children were hearing a story. He was introduced to the children, commented on the story, and walked around the perimeter of the classroom. He then departed, having stayed a total of approximately two minutes.[88] Following Sam Stone's visit, researchers interviewed the children four times over a four-week period. In the last three interviews, children were provided with "erroneous suggestions . . . that Sam Stone had ripped a book [and] . . . soiled a teddy bear."[89] For example, in the second interview, interviewers asked the children "Did Sam Stone rip the book with his hands, or did he use scissors?"[90]

Approximately ten weeks after Sam Stone's visit, a new interviewer questioned the children. The interviewer first asked a "free-narrative" question: "Remember the day that Sam Stone came to your classroom? Well, I wasn't there that day, and I'd like you to tell me everything that happened when he visited."[91] If the child did not specifically refer to a book being ripped or a teddy bear being soiled, she was asked "probe" questions: "I heard something about a book. Do you know anything about that?" and "I heard something about a teddy bear. Do you know anything about that?"[92] Forty-six percent of the three- and four-year-old children spontaneously reported that Sam Stone had performed one or both misdeeds in response to the free narrative question; seventy-two percent did so in response to probe questions.[93]

> What was most surprising about these children's reports was the number of false perceptual details, as well as nonverbal gestures, that they provided to embellish their stories of these nonevents. For example, children used their hands to show how Sam had purportedly thrown the teddy bear up in the air; some children reported seeing Sam in the playground, on his way to the store to buy chocolate ice cream, or in the bathroom soaking the teddy bear in water before smearing it with a crayon.[94]

---

87 *See* Leichtman & Ceci, *supra* note 34, at 570.

88 *See id.*

89 *Id.* at 571.

90 *Id.* app. B at 577 (internal quotation marks omitted).

91 *Id.* app. C at 578 (internal quotation marks omitted).

92 *Id.* (internal quotation marks omitted).

93 *See id.* at 572-73.

94 CECI & BRUCK, *supra* note 36, at 131-32.

B. Bruck, Ceci, Francoeur, and Barr's Inoculation Study

In Bruck, Ceci, Francoeur, and Barr's Inoculation Study, a pediatrician gave four- and five-year-old children a routine medical examination.[95] After the examination, a research assistant greeted the children and spoke to them about a poster on the wall for several minutes. The research assistant stayed during the pediatrician's administration of the oral vaccine and the inoculation and then took the child to another room where she gave them treats and read them a story.[96]

Approximately eleven months after their visit to the pediatrician, researchers interviewed the children four times over a two-week period.[97] In the first three interviews, researchers gave the children false information about their visit. The interviewer minimized how much the inoculation had hurt and how much the children had cried.[98] In addition, the interviewer told the children that the research assistant had given them their oral vaccine and inoculation, and that the pediatrician had shown them the poster, given them the treats, and read them the story.[99] In the fourth interview, researchers asked the children to recall everything that happened on their visit to the pediatrician's and directly asked who had performed the various actions during their visit (if the children had not already volunteered such information).[100]

In the fourth interview, the children reported significantly less pain and crying than a control group of children. About thirty percent to forty percent of the children falsely reported that the research assistant had given them their shot, the oral vaccine, and the checkup, and that the pediatrician had shown them the poster, given them treats, and read them a story.[101] The authors concluded, "[t]hese results challenge the view that suggestibility effects are confined to peripheral, neutral, and non-meaningful events."[102]

C. Ceci, Crotteau Huffman, and Smith's Mousetrap Study

In Ceci, Crotteau Huffman, and Smith's Mousetrap Study, researchers interviewed preschool children about various events, only some of which had occurred, seven to ten times over a period of ten

95 *See* Bruck et al., *I Hardly Cried*, *supra* note 34, at 195.
96 *See id.*
97 *See id.* at 198-99.
98 *See id.* at 200-01.
99 *See id.*
100 *See id.*
101 *See id.* at 203-04.
102 *Id.* at 207.

weeks.[103] One of the fictitious events concerned getting one's hand caught in a mousetrap and having to go to the hospital. The experimenter held cards on which the events were written and told the child that only some of the events had occurred and that the child should "think real hard" and decide whether each event had really happened or not.[104] At the end of ten weeks, a new interviewer asked the children whether the events had ever occurred.[105]

Fifty-eight percent of the children produced false narratives to at least one of the fictitious events, and twenty-five percent falsely affirmed that most of them had occurred.[106] Many children were able to provide compelling narrative accounts of the nonexistent events. For example:

> "My daddy, mommy, and my brother [took me to the hospital] in our van. . . . The hospital gave me a little bandage, and it was right here [pointing to index finger] . . . . I was looking and then I didn't see what I was doing and it [finger] got in there somehow. . . . The mousetrap was in our house because there's a mouse in our house . . . . The mousetrap is down in the basement, next to the firewood . . . . I was playing a game called 'operation' and then I went downstairs and said to Dad, 'I want to eat lunch,' and then it got stuck in the mousetrap . . . . My daddy was down in the basement collecting firewood . . . . [My brother] pushed me [into the mousetrap]; he grabbed *Blow Torch* [an action figure]. It happened yesterday. The mouse was in my house yesterday. I caught my finger in it yesterday. I went to the hospital yesterday."[107]

D. Bruck, Hembrooke, and Ceci's Monkey-Thief Study

In Bruck, Hembrooke, and Ceci's Monkey-Thief Study, researchers interviewed sixteen preschool children on five occasions about four events: two true events and two false events.[108] One of each type of event was a positive event, and one was a negative event. The false-positive event involved helping a woman find her lost monkey, whereas the false negative involved witnessing a man come to the day-care and steal food. In the first interview, the researcher simply asked the children whether the events had occurred. In the second and third interviews, the interviewers used a combination of suggestive

---

103 *See* Stephen J. Ceci et al., *Repeatedly Thinking About a Non-Event: Source Misattributions Among Preschoolers*, 3 CONSCIOUSNESS & COGNITION 388, 394 (1994).

104 *Id.* at 394, 395 (internal quotation marks omitted).

105 *See id.* at 395.

106 *See* CECI & BRUCK, *supra* note 36, at 219.

107 Stephen J. Ceci et al., *The Possible Role of Souree Misattributions in the Creation of False Beliefs Among Preschoolers*, 42 INT'L J. CLINICAL & EXPERIMENTAL HYPNOSIS 304, 306-07 (1994) (omissions and alterations in original) (quoting one of the children in the study).

108 *See* Maggie Bruck et al., *Children's Reports of Pleasant and Unpleasant Events*, *in* RECOLLECTIONS OF TRAUMA 199, 203 (J. Don Read & D. Stephen Lindsay eds., 1997).

techniques that included "peer pressure, visualization techniques, repeating misinformation, and selective reinforcement."[109] If the children stated that an event had occurred, the interviewer asked open-ended and closed-ended questions about the event. If the children denied that the event had occurred, the interviewer asked them to pretend that it had and asked the same questions. On the fourth interview, the researcher asked the children to tell their stories to a puppet. Again, if the children denied that an event had occurred, the researcher asked them to pretend. On the fifth interview, a new interviewer asked an open-ended question about the events (e.g., "I heard something about a lost monkey. Do you know anything about that?").[110] The study found that "[b]y the third interview, most children had assented to all true and false events. This pattern continued to the end of the experiment."[111]

These studies undercut sanguine assumptions that children are not unduly suggestible. In each of these studies, a substantial number of children falsely affirmed that nonexistent events had occurred. These false reports often occurred spontaneously, in response to a request for free narrative. Moreover, children frequently elaborated on their false reports, even going beyond the information previous interviewers had suggested. Finally, the false reports often concerned events in which the children both participated and were harmed. The results thus challenge the shibboleths of previous research on children's suggestibility: false reports occur rarely and only in response to highly misleading questions;[112] false reports tend to be unelaborated, single-word responses;[113] false reports are unlikely when the child is reporting a negative event that involves the child's body.[114]

---

109 *Id.* at 204.

110 *Id.*

111 Bruck et al., *supra* note 30, at 143.

112 *See* Goodman & Reed, *supra* note 19, at 328 ("While children appeared more suggestible than adults, suggested information was unlikely to appear in their free recall of the event.").

113 *See* Gail S. Goodman & Christine Aman, *Children's Use of Anatomically Detailed Dolls To Recount an Event*, 61 CHILD DEV. 1859, 1869 (1990) ("[T]he children's errors were largely nods of the head. The children never provided spontaneous elaborations that would indicate that sexual abuse occurred."); Goodman et al., *Sexual and Physical Abuse*, *supra* note 19, at 12 ("[W]hen the children conformed to the suggestive questions, it was typically with a hesitant yes or no, without further incorrect elaboration. The children were more likely to elaborate on their correct responses."); Saywitz et al., *supra* note 59, at 687 ("Of the [three] children in the nongenital condition who made the three commission errors [and falsely reported genital and/or anal touch], two were unable to provide any detail.").

114 *See* Goodman et al., *supra* note 23, at 278-79 ("By the age of 4 years, most children we have tested are surprisingly resistant to abuse suggestions. . . . The answer may lie in the fact that child abuse involves actions directed against a child's body, actions that violate their concerns.").

On the other hand, the new wave studies establish only that researchers can produce false allegations and do not enable others to estimate how often such allegations are occurring under current interviewing practices. To make such a judgment, one must understand how investigators actually conduct these interviews in the real world. Such an understanding leads to the conclusion that the new wave research may overstate children's suggestibility in actual practice. Moreover, the new wave ignores a number of important variables in their criticism of interviewing practices. These variables decrease the likelihood of false allegations of sexual abuse and in some cases justify the use of interviewing practices the new wave criticizes.

III

THE REAL WORLD OF SEXUAL ABUSE INVESTIGATIONS

A. The Representativeness of Interviews the New Wave Reviews

To determine the practical relevance of new-wave research, one needs to know the extent to which interviewers do in fact use the suggestive techniques the new wave examines. Noting that the interviewing strategies have negative effects or even that some interviewers have used the techniques is not sufficient. In many situations, legislators and courts must make decisions based on the way investigators typically conduct interviews. If interviews are usually suggestive, then one fairly can presume that the interviews in a particular case were suggestive. A presumption that interviews are suggestive affects legislative decision making regarding the admissibility of children's statements and judicial decision making regarding whether to admit evidence in particular cases. Moreover, such a presumption justifies the use of expert testimony to review the results of the new wave's research findings.

Presumptions may reflect the relative weights one assigns to the two types of error: (1) an erroneous assumption that the interviewing technique was not suggestive and (2) an erroneous assumption that the interviewing was suggestive. However, unless one believes that any risk of one type of error trumps the other type of error, no matter what its magnitude, one must establish what interviews are actually like in order to make an informed value judgment regarding the suggestiveness of child interviewing.

Information regarding what typically occurs in interviews might seem irrelevant with respect to individual cases. Obviously, if one knows whether a particular child was interviewed with suggestive techniques, then one need not ask what most interviews are like. However, how interviewers conducted interviews is largely unknown in many, if not most, cases. Although many jurisdictions require videotaping or taping of investigatory interviews, most do not. Further-

more, it would be impractical to impose a requirement that individuals record the first contact with the child giving rise to a suspicion of abuse because such contact arises between children and parents or teachers, rather than during a formal abuse investigation. Documentation, when it occurs, often fails to provide verbatim information regarding the child's disclosures, and interviewers may be unable to recall the exact wording of their questions. In sum, to make judgments in individual cases, courts often must make assumptions about how interviewers typically interview children. Furthermore, experts who testify about suggestive methods of questioning cannot be sure whether the research they discuss is relevant to the particular case, especially if they take Ceci and Bruck's advice that they should learn very little about the case (save the child's age) to remain impartial.[115]

At first glance, the new wave makes claims about the nature of interviewing in general when criticizing various interviewing techniques as unduly suggestive. Discussing the ecological validity of their research, Ceci and Bruck argued that

> [t]he major differences between suggestive interviews in laboratory studies and suggestive interviews in actual cases is that the former are generally less intense and contain fewer suggestive elements than the latter. . . . This leads to the conclusion that if experiments were more like real-life cases we would elicit many more false reports from children than we have done to date . . . .[116]

Differences between the new wave and Goodman and her colleagues may derive more from differing assumptions about what inter-

---

115 *See* CECI & BRUCK, *supra* note 36, at 276 n.1. Discussing her testimony in the *Sterling* case, Bruck asked the defense attorney before trial for a "brief outline of the facts of the case as well as some material on the interviewing procedures used with the children so that I could he sure that the suggestibility of young children was a key issue in this case." Bruck, *supra* note 44, at 99. On cross-examination in her testimony, she responded negatively to the prosecutor's question as to whether she had done "any work on this case in terms of looking at the techniques used." Testimony of Maggie Bruck at 5256, The Queen v. Sterling, Q.B.J. No. 74 (Sask.) (1994).

116 CECI & BRUCK, *supra* note 36, at 299; *see also* Bruck & Ceci, *supra* note 30, at 207 ("We have chosen to include in our studies procedures to measure the risks of various interviewing techniques that are *commonly* used in investigative and therapeutic arenas." (emphasis added)); *id.* at 211 (arguing that "adults make frequent misleading suggestions" in "a climate more *typical* of actual sex abuse investigations" (emphasis added)); Bruck et al., *supra* note 108, at 199 (noting that social scientists have turned their attention "to examining the accuracy of children's testimony under a range of conditions that are *characteristic* of those that bring children to court" (emphasis added)). Other researchers have made similar claims. *See* Debra Ann Poole & Lawrence T. White, *Tell Me Again and Again: Stability and Change in the Repeated Testimonies of Children and Adults*, *in* MEMORY AND TESTIMONY IN THE CHILD WITNESS 24, 32 (Maria S. Zaragoza et al. eds., 1995) ("[F]ew investigators have simulated the intensity of leading and misleading information to which the typical witness is exposed. Notable exceptions are two recent studies by Ceci and his colleagues." (citing the Sam Stone Study and the Inoculation Study)).

views are like than from differing beliefs about children's vulnerability. Saywitz, Goodman, and their colleagues believed that their interviews of girls examined by pediatricians[117] were "ecologically valid" in that they replicated actual child abuse interviews, whereas Bruck dismissed the study as "meaningless" because the interviews were totally unlike those in the real world.[118] Goodman, Rudy, Bottoms, and Aman argued that anyone who asked questions as leading as those in her own studies "would likely face severe criticism from the accused that the child was led into making a false accusation. Child abuse charges have often been dismissed by judges on this ground."[119] In contrast, Ceci and Bruck criticized what may have been Goodman's most extreme manipulation: the study in which the researchers created an "atmosphere of accusation" by telling the children interviewers would question them about an important event and by saying things like "Are you afraid to tell?" and "You'll feel better once you've told." Ceci and Bruck claimed that "the *typical* forensic case would have involved *multiple* prior attempts to create an 'atmosphere of accusation,' not just a single one several years after an event."[120]

When pressed, however, the new wave hedges on the position that the typical interview is in fact anything like those used in their research. In the *Kelly* case, Bruck testified on direct examination that even the most careful of interviewers will lapse into suggestiveness,[121] and in the *Sterling* case, Bruck stated on cross examination that her descriptions of suggestive techniques were "typical."[122] As the subsequent questioning made clear, however, she based her views of what constitutes typical interviewing on transcripts that professionals, primarily defense attorneys, had sent to her over the years.[123]

---

117 *See supra* text accompanying notes 60-69.

118 *Compare* Saywitz et al., *supra* note 59, at 690-91 (noting that the study "attempted to attain greater ecological validity than heretofore achieved in research on children's testimony" and that it fell short only to the extent that it "lacked the urgency of a clinical evaluation or courtroom proceeding"), *with* Testimony of Maggie Bruck at 15,458-59, State v. Kelly, No. 933SC676 (N.C. Super. Ct. 1992) ("As I said about the Goodman and, ah, Saywitz study. It's a meaningless study. Those kids were questioned in totally unrealistic ways in terms of what goes on in sexual abuse cases.").

119 Goodman et al., *supra* note 23, at 258.

120 Ceci & Bruck, *supra* note 9, at 421 (first emphasis added).

121 *See* Testimony of Maggie Bruck at 15,332, *Kelly*, No. 933SC676 ("Q: In your experience and based upon the scientific studies that you reviewed, . . . do even the most careful interviewers resort or lapse into the use of suggestive, . . . questioning techniques? A: Yes, they do.").

122 *See* Testimony of Maggie Bruck at 5256, The Queen v. Sterling, Q.B.J. No. 74 (Sask.) (1994) ("Q: So you are generalizing in terms of techniques. Is that correct? A: Yes. Q: Okay. Now you talk about what typically occurs in an investigative interview. Is that correct? A: That's correct.").

123 *See id.* at 5257-58.

These transcripts also provide the basis for Ceci and Bruck's book, in which they warned the reader that "the materials we have reviewed may not be representative of many of the interviews carried out with children in forensic or therapeutic situations."[124] These transcripts came to the authors' attention "because they contain[ed] components that might be considered to be suggestive";[125] therefore, one can fairly assume that nonsuggestive interviews were underrepresented.

Moreover, the review of the transcripts was not "scientific,"[126] and the suggestive elements were not quantified.[127] Bruck looked through the transcripts informally to determine if suggestive techniques were present.[128] At best, the review process made it impossible to determine if those interviews that do contain, for example, stereotype induction, do so in any substantial way. At worst, one wonders whether the authors' own expectations exerted an influence on their interpretation. As Ceci and Bruck warned:

> Expectations and biases affect how situations are encoded and subsequently remembered. Generally, expectancy-consistent results are more likely to be remembered: The number of confirming cases are overestimated, and these confirming cases are more easily recalled. Prior expectations (or biases) may also work on incongruent information in such a way as to transform it so that it fits into one's existing beliefs.[129]

The problems of unsystematic review of interviews are apparent when one reads that in 1994 Ceci found "potentially suggestive and stereotype inducing" methods in one-third of the cases he reviewed,[130] whereas in 1995 Ceci and his colleagues found "highly improper interviewing techniques" in the "vast majority of cases" they reviewed.[131] One cannot determine whether interviews grew worse, the interviews the authors *received* grew worse, or the criteria for judging the interviews changed.

Ultimately, the new wave cited some systematic research on interviewing, discussed below, but simultaneously refused to "endorse any specific prevalence rate of poor interviews." Instead, the new wave

---

124 CECI & BRUCK, *supra* note 36, at 82.

125 *Id.*

126 Testimony of Maggie Bruck at 5259, *Sterling*, Q.B.J. No. 74 (Sask.) ("[T]his isn't a scientific study of investigative techniques.").

127 *See* CECI & BRUCK, *supra* note 36, at 82.

128 *See* Testimony of Maggie Bruck at 5260-61, *Sterling*, Q.B.J. No. 74 (Sask.) ("I haven't really monitored myself about how I do this, but I don't—I don't review interviews for scientific purposes. I review the interviews for general knowledge about what this interview might have been like, but it's not to do a scientific study on.").

129 CECI & BRUCK, *supra* note 36, at 88 (citations omitted).

130 Ceci, *supra* note 32, at 26.

131 Ceci et al., *supra* note 33, at 501.

took the position that there are "many" improper interviews, and "[w]hether these represent a substantial portion or only a tiny portion of all cases is anyone's guess."[132] Ceci most recently has expressed the "hunch . . . that the majority of interviews done with kids by front-line workers, child-protective service, law enforcement, therapists, pediatricians, are well-done."[133] The position that "many" interviews are improperly suggestive emphasizes the possibility rather than the relative probability of false positives. If the mere existence of bad interviews presents a sufficient ground for policymaking, the implicit value judgment is that no false positives are tolerable.

B. Day-Care Cases Versus Typical Abuse Cases

Questions regarding the representativeness of the case studies that the new wave used to illustrate the dangers of interviewing techniques relate to concerns over the representativeness of the interviewing techniques the new wave criticized. Ceci and Bruck warned that "most of the actual sexual abuse cases that we describe are day-care cases in which some of the children make allegations of ritualistic abuse at the hands of their caregivers."[134] Besides the Salem Witch trials, the authors described six contemporary cases of sexual abuse, four of which involved allegations by large numbers of preschool children in day care.[135] They acknowledged that these cases "represent only a small subset of the actual sexual abuse cases."[136] Nevertheless, they used these cases as "'windows'" through which one can understand the problems of child sexual abuse allegations.[137]

Ceci and Bruck defended their emphasis on day-care cases (which involve multiple preschool children alleging ritualistic abuse) on at least three grounds. First, they argued that "although these cases represent only a small proportion of sexual abuse complaints, in absolute numbers they involve a large number of children (in the Mc-

---

132 *Id.*; *see id.* ("[W]e wonder whether any knowledgeable individual believes that poor interviews do not exist. Again, we answer our own question by pointing out that Myers (1995) himself, despite his posturing to the contrary, finally agrees with our claims: 'I concede that highly improper interviews occur too often.'").

133 *Nightline* (ABC television broadcast, Nov. 14, 1996), *available in* LEXIS, ABC News Transcripts. Compare this view to Ralph Underwager, who is quoted in the same broadcast: "I believe the great majority of the questioning of children that is done in this country is highly coercive, highly suggestive, leading, and produces inaccurate information." *Id.*

134 CECI & BRUCK, *supra* note 36, at x.

135 *See id.* at 9 (describing *Little Rascals* with twelve child witnesses); *id.* at 11-13 (describing *Michaels* with nineteen child witnesses, and *Finje/Old Cutler* with a "large number of preschool children"); *id.* at 15 (describing *Fuster/Country Walk* with five child witnesses). The other two cases involved a rape of a seven-year-old girl, allegations the authors believed were true, and a murder case involving a nine-year-old witness, allegations the authors suspected were untrue. *See id.*

136 *Id.*

137 *See* Ceci & Bruck, *supra* note 9, at 403.

Martin case, for instance, interviewers under contract to the State of California alleged the abuse of 369 children)"[138] and "in other day-care cases the number of allegations is also quite large."[139] However, the *McMartin* case is hardly representative of day-care cases. As Finkelhor's study of day-care abuse emphasized:

> One clear-cut way in which the McMartin Preschool case was atypical of day-care abuse was in the enormous number of children involved. Investigators estimated the number of victimized children at more than 300, spanning a period of at least 10 years. By contrast, the majority of other day-care abuse cases involved the substantiated abuse of only one or two children.[140]

Moreover, if our concern is that a large number of defendants are falsely accused, we should focus on the number of cases in which multiple victims testified. Eighty-five percent of criminal sexual abuse cases involve only a single victim.[141]

Second, Ceci and Brnck asserted that "day-care cases are relevant to the more general testimonial issues found in many nonday-care cases (i.e., repeated suggestive questioning, interviewer stereotypes, failure to test alternative hypotheses)."[142] However, consider how the dynamics of an abuse investigation change when one compares multivictim cases with single-victim cases. Interviewers who are confident that the children have suffered abuse are more likely to question extensively a child in a multivictim case than in a single-victim case. Likewise, interviewers are more likely to confront a child with the statements of other children alleging abuse in multivictim cases. Interviewers in multivictim cases also are more likely to "assure" a child at the beginning of questioning that the accused is a bad person who has hurt children and who has been put in jail. In a single-victim case, the investigator will question the child because a specific suspicion exists that the child has suffered abuse. If the child is the only victim, contamination by other victims is obviously impossible. Until the

---

138 Stephen J. Ceci & Maggie Bruck, *Child Witnesses: Translating Research Into Policy*, SOC. POL'Y REP., Fall 1993, at 1, 26 n.3 (citation omitted).

139 CECI & BRUCK, *supra* note 36, at xi.

140 DAVID FINKELHOR ET AL., NURSERY CRIMES: SEXUAL ABUSE IN DAY CARE 70 (1988). Nine children ultimately testified at the *McMartin* trial. *See* NANCY WALKER PERRY & LAWRENCE S. WRIGHTSMAN, THE CHILD WITNESS 8 (1991).

141 *See* SMITH & ELSTEIN, *supra* note 15, at 82 ("In over four-fifths of the cases (85%), the defendant was charged with sexually abusing a single victim. In an additional 11% of the cases, two victims were named in the indictment. Far less often (5% of the cases), three or more victims were named." (footnotes and citations omitted)); *see also id.* at 99 (noting that in 22% of cases in which a defendant was charged with abuse of one victim there were allegations of abuse against other victims which were not charged).

142 Ceci & Bruck, *supra* note 138, at 26 n.3; *cf.* CECI & BRUCK, *supra* note 36, at xi ("[A]ll of the arguments we make for and against the reliability of children's testimony in sexual abuse cases apply equally to non-sexual-abuse contexts . . . .").

child acknowledges some abuse, it is unlikely that the accused will be labelled a criminal.

Other differences exist between day-care cases and typical sexual abuse cases that make the likelihood of suggestive questioning greater in the day-care cases. Most sexual abuse cases involve abuse by someone close to the child—typically a family member or friend of the family.[143] Parents are unlikely to pursue the hypothesis that a spouse or a brother has abused their child. Furthermore, interviewers are not likely to paint negative stereotypes of those with whom the child may wish to maintain an ongoing relationship. Surely, some cases arise in which negative stereotyping occurs,[144] but in most cases those close to the child are not eager to believe that someone has abused the child.[145] The median age of a sexual abuse victim in criminal court is thirteen years of age,[146] while the day-care cases predominantly involved preschool children. As I will argue below, these and other differences also mean that children will be less suggestible in the truly "typical" case of sexual abuse.

The final justification the new wave gave for emphasizing day-care cases is that, "because of their visibility, day-care cases are often more extensively documented."[147] Cases journalists and legal scholars have brought to public attention receive the best documentation and tend to be those which "cast doubt on the accuracy of children's statements."[148] A scientific study of children's suggestibility that focuses on highly visible cases, however, is much like a report on airline safety that focuses on air disasters. Coverage of air disasters likely leads the public to overstate vastly the dangers of air travel;[149] coverage of dubi-

---

143 *See* CECI & BRUCK, *supra* note 36, at 30 ("[I]t appears that most of the reported cases of child abuse involve intrafamilial abuse (broadly defined) and that only a small minority involve strangers."); SMITH & ELSTEIN, *supra* note 15, at 86 (studying the relationship between defendant and victim in criminal child abuse cases and concluding that "[o]nly 6% of the defendants were strangers to their victims" and "[t]he most common relationship was that of a parent, or a parental figure"). Of course, virtually all alleged abusers in dependency and family court are family members because these courts only hear cases in which family members are allegedly responsible for abuse, either because they themselves are abusers or because they allowed someone else to abuse their child.

144 Custody battles in which ex-spouses are accused of abuse present the most likely occasion for stereotype induction. It is unclear, however, whether one needs to warn judges who hear these cases about the dangers of influence because family courts generally are skeptical of sexual abuse claims. *See* Meredith Sherman Fahn, *Allegations of Child Sexual Abuse in Custody Disputes: Getting to the Truth of the Matter*, 25 FAM. L.Q. 193, 194-95 (1991).

145 *See infra* text accompanying notes 288-96 (arguing that mothers are often unsupportive of their children's allegations of abuse).

146 *See* SMITH & ELSTEIN, *supra* note 15, at 83 tbl.IV-1.

147 Ceci & Bruck, *supra* note 138, at 26 n.3; *see* CECI & BRUCK, *supra* note 36, at xi.

148 CECI & BRUCK, *supra* note 36, at x.

149 *See* Paul Slovic et al., *Facts Versus Fears: Understanding Perceived Risk*, *in* JUDGMENT UNDER UNCERTAINTY: HEURISTICS AND BIASES 463, 467-68 (Daniel Kahneman et al. eds., 1982).

ous child sex abuse cases surely will have a similar effect. At the same time, the media's emphasis on bizarre, facially implausible allegations falsely assures the public that society will not miss true cases of abuse and creates the perception that reining in interviewers poses no potential risks. Analogies to the Salem Witch trials[150] have the same effect. The public need not fear that increased skepticism will give witches free rein to practice their craft.

C. Leading Questions in Practice and in Research

In their most recent paper, Bruck, Ceci, and Hembrooke discuss a number of "suggestive interviewing techniques," including leading questions, stereotype inducement, selective reinforcement, guided imagery, and peer pressure.[151] To "dispute [the] claim" that "the interviewing conditions used in the suggestibility studies are rarely used by professionals,"[152] the authors cite several studies on real-world interviewing that allegedly support the ecological validity of their research.

The studies do not indicate, however, that stereotype inducement or selective reinforcement is common. The studies do not measure these tactics. Nor do they document extensive use of peer pressure or guided imagery. Indeed, the research suggests that the latter two sources of suggestibility are *uncommon.* Warren and colleagues examined the extent to which an interviewer told a child that another person had said the child was abused—a method which could constitute peer (or adult) pressure. Interviewers employed this tactic three times out of forty-two interviews.[153] Boat and Everson examined the

---

150 *See* CECI & BRUCK, *supra* note 36, at 8-9.

151 Bruck et al., *supra* note 30, at 140-41. The authors also discuss the use of anatomically correct dolls. *See id.* at 140. Whether one can characterize their use per se as evincing "interviewer bias" is beyond the scope of this paper; suffice it to say that several professional organizations have taken the position that the use of dolls as an adjunct to interviewing is appropriate. *See* AMERICAN PROF'L SOC'Y ON THE ABUSE OF CHILDREN, PRACTICE GUIDELINES: USE OF ANATOMICAL DOLLS IN CHILD SEXUAL ABUSE ASSESSMENTS 2 (1995) ("When used by a knowledgeable and experienced professional, anatomical dolls can be an effective tool to aid in interviewing children."); Ronald E. Fox, *Proceedings of the American Psychological Association, Incorporated, for the Year 1990: Minutes of the Annual Meeting of the Council of Representatives,* 46 AM. PSYCHOL. 689, 722 (1991) (issuing a statement on the use of anatomically detailed dolls in forensic evaluations adopted by APA's Council of Representatives and approving "doll-centered assessment of children when used as part of a psychological evaluation and interpreted by experienced and competent examiners"). *But cf.* Gerald P. Koocher et al., *Psychological Science and the Use of Anatomically Detailed Dolls in Child Sexual-Abuse Assessments,* 118 PSYCHOL. BULL. 199, 218 (1995) (reporting for the Anatomical Doll Working Group financed by the APA and concluding that although dolls "can still provide a useful communication tool in the hands of a trained professional interviewer," the APA should "reconsider whether valid 'doll-centered assessment' techniques exist").

152 Bruck et al., *supra* note 30, at 143.

153 *See* Amye R. Warren et al., *"It Sounds Good in Theory, But . . .": Do Investigative Interviewers Follow Guidelines Based on Memory Research?,* 1 CHILD MALTREATMENT 231, 238-39 (1996).

extent to which interviewers using dolls asked the child to show how abuse might have occurred[154]—a method which could be similar to guided imagery.[155] Interviewers *never* employed this tactic in ninety-seven interviews.[156]

The research, however, *does* document infrequent use of open-ended questions with alleged child abuse victims. Bruck and her colleagues summarized the observational research on real-world interviews and concluded that "interviewers mainly relied on specific or leading questions; several times during the interviews, they introduced information that the children had not volunteered, and they frequently repeated that new information in the course of a single interview."[157] The authors rely primarily on work by Amye Warren and colleagues, and Michael Lamb and colleagues.[158]

Warren and her colleagues examined transcripts of interviews conducted in the late 1980s and early 1990s by child protective services workers in a southern state.[159] They found that nearly ninety percent of the questions asked by interviewers constituted "specific" questions because they did not require a narrative response.[160] Noting that "questions containing previously undisclosed information may be considered leading questions,"[161] the authors found that interviewers introduced, on average, seven pieces of new information per interview[162] and, again on average, repeated the new information once during the interview.[163] Examples of interviewers providing new information included stating the following: where the child currently lives, names of people the child knows, what the child's mother has

---

154 *See* Barbara W. Boat & Mark D. Everson, *Concerning Practices of Interviewers When Using Anatomical Dolls in Child Protective Services Investigations*, 1 CHILD MALTREATMENT 96 (1996).

155 *See* Mary Ann Foley et al., *Developmental Comparisons of the Ability To Discriminate Between Memories for Symbolic Play Enactments*, 30 DEVELOPMENTAL PSYCHOL. 206, 208-09 (1994) (finding that three-year-olds are more likely than five-year-olds to believe mistakenly that they had played with an object that they had only pretended to play with).

156 *See* Boat & Everson, *supra* note 154, at 101 & tbl.4.

157 Bruck et al., *supra* note 30, at 141.

158 *See* CECI & BRUCK, *supra* note 36, at 82. They stated that

> some recent studies reveal the pervasiveness of some interviewing styles. Warren and her colleagues (see McGough and Warren, 1994) have analyzed the child sexual abuse investigative interviews conducted by Child Protective Services professionals in the state of Tennessee. These interviewers spent little if any time asking children open-ended questions; 90% of all questions were highly specific, requiring one-word answers (see Lamb et al., in press, for similar results for trained Israeli "youth" investigators).

Id.

159 *See* Warren et al., *supra* note 153, at 232.

160 *See id.* at 236.

161 *Id.* at 241.

162 *See id.* Approximately 94% of the interviewers in the Warren sample introduced new information in their interview. *See id.*

163 *See id.* at 242.

done, and the fact that the child has spoken to the interviewer before.[164] As a worst-case example of providing new abuse details, Warren and colleagues quote an interviewer who reminded the child what she *previously had told the interviewer about abuse.*[165]

Warren and her colleagues cited Lamb's findings as consistent with their own.[166] Lamb and his colleagues examined Israeli youth investigators' interviews with alleged sexual abuse victims and found that only two percent of the utterances were "invitations," defined as statements that invite "an open-ended response from the child."[167] Lamb also found that about twenty-five percent of the investigators' utterances were leading.[168] Lamb's definition of "leading," however, included any statements that "focus the child's attention on details or aspects of the account that the child has not previously mentioned, *but do not imply that a particular response is expected.*"[169] The study classified questions that imply a desired response or that assume details that the child had not provided as "suggestive."[170] Lamb found that investigators posed such suggestive questions approximately nine percent of the time.[171]

Lamb's research group obtained similar results from three other interview samples: one conducted by sheriff's investigators in a small

---

164 *See id.* at 241-42.

165 *See id.* at 242.

166 *See id.* at 237 (noting that Lamb and colleagues obtained "consistent results," because in Lamb's work, "[f]ewer than 10% of the interviewers' utterances were invitational, defined as encouraging an open-ended, narrative response from children"). At times, Warren and colleagues' findings make interviewers look worse than those in Lamb's research because of the way Warren and colleagues present their data. Rather than calculate the percentage of questions that introduced new information, Warren and colleagues calculated the percentage of interviewers who *ever* introduced new information. *See id.* at 241-42. Hence, approximately 94% of the interviews introduced new information, and 67% repeated new information at some point in the interview. *See id.* These are impressive numbers, but consistent with the proposition that proportionally very few questions impart new information to the child.

167 Michael E. Lamb et al., *Effects of Investigative Utterance Types on Israeli Children's Responses*, 19 INT'L J. BEHAV. DEV. 627, 631, 633 (1996). Lamb's research identifies the Israeli youth investigators studied as "specially trained," implying that they might outperform the typical American investigator. Warren et al., *supra* note 153, at 237; *see also* CECI & BRUCK, *supra* note 36, at 82 (noting that Israeli youth investigators are "trained"). Most, however, received training as probation officers rather than as forensic investigators. *See* Kathleen J. Sternberg et al., *Child Sexual Abuse Investigations in Israel: Evaluating Innovative Practices, in* INTERNATIONAL PERSPECTIVES ON CHILD ABUSE AND CHILDREN'S TESTIMONY 62, 69 (Bette L. Bottoms & Gail S. Goodman eds., 1996) ("Youth investigators are not adequately trained in forensic investigation and must therefore learn the necessary skills 'on the job.'"). Furthermore, the *least* experienced investigators were assigned the sexual abuse cases. *See id.*

168 *See* Lamb et al., *supra* note 167, at 633 tbl.1.

169 *Id.* at 631 (emphasis added).

170 *Id.*

171 *See id.* at 633 tbl.1.

southern town,[172] one conducted by protective service workers in a large southeastern state,[173] and one conducted by "two expert and experienced forensic psychologists."[174] Across all three studies, approximately ten percent of interviewers' questions are "suggestive," and an average interview contained from five to ten suggestive statements.[175] The data do not reveal the extent to which the suggestive questions elicited details of the alleged abuse.[176]

In sum, the limited observational research on real-world interviews demonstrates that interviewers ask few open-ended questions, many specific questions, and some leading questions.[177] The proportion of leading questions depends on how narrowly one defines "leading." How does this compare with the new-wave research? While the

---

172 *See* Kathleen J. Sternberg et al., *The Relation Between Investigative Utterance Types and the Informativeness of Child Witnesses*, 17 J. APPLIED DEVELOPMENTAL PSYCHOL. 439, 442, 447-49 (1996).

173 *See* Michael E. Lamb et al., *Investigative Interviews of Alleged Sexual Abuse Victims with and Without Anatomical Dolls*, 20 CHILD ABUSE & NEGLECT 1251, 1252, 1256 (1996).

174 Irit Hershkowitz et al., *The Relationships Among Interviewer Utterance Type, CBCA Scores and the Richness of Children's Responses*, 2 LEGAL & CRIMINOLOGICAL PSYCHOL. 169, 171 (1997); *see id.* at 173-74.

175 *See* Lamb et al., *supra* note 173, at 1255 tbl.1 (finding 7.2% suggestive questions with an average of 4.9 utterances when no doll was used and 8.2% with an average of 8 utterances when a doll was used); Sternberg et al., *supra* note 172, at 446 tbl.2 (finding 8.7% suggestive questions when multiple incidents of abuse alleged and 9.9% when single incident alleged).

176 The other studies Bruck and her colleagues cite exemplify rather than address the problems with defining leading and suggestive questioning. Bull and Cherryman, for example, asked evaluators to listen to police interrogations and to rate how often investigators asked "leading questions," yet they never defined the term "leading question." On average, the evaluators found that leading questions were "often present." RAY BULL & JULIE CHERRYMAN, HELPING TO IDENTIFY SKILLS GAPS IN SPECIALIST INVESTIGATIVE INTERVIEWING 20 (1995) (internal quotation marks omitted). The fact that the report examined interviews with *suspects*, most of them *adults*, colors the relevance of the report for assessing interviews with children. The study by Yuille and his colleagues is unpublished, and unfortunately the published accounts do not define the "inappropriate interviewing" which rendered "meaningless any assessment of the child's account based on the interview alone," a problem the authors identified among a fourth of the investigative interviews they examined. John C. Yuille et al., *The Nature of Allegations of Child Sexual Abuse*, *in* TRUE AND FALSE ALLEGATIONS OF CHILD SEXUAL ABUSE 21, 35 (Tara Ney ed., 1995).

177 Several factors may limit the generalizability of these findings. For example, interviewers might behave differently when questioning younger children. Although half of the interviewees in Warren and colleagues' research were six years of age or younger, *see* Warren et al., *supra* note 153, at 232, only a small proportion of interviewees in Lamb and colleagues' samples were of preschool age, *see* Lamb et al., *supra* note 167, at 635 ("Because we limited this study to children between 5 and 11 years of age, we do not know whether similar results would have been obtained had the study been focused on preschool-aged children."). Moreover, clinicians who accept special referrals for sexual abuse evaluation might differ from front-line child protective service workers and police officers. Clinicians may see reticent, troubled children more often, thereby prompting them to ask more leading questions. *But see* Hershkowitz et al., *supra* note 174, at 171-72 (characterizing 11% of questions as suggestive in a sample of interviews by "two expert and experienced forensic psychologists" who conducted interviews "at the request of legal, judicial and criminal justice agencies").

new wave fails formally to define "leading," Bruck and Ceci provided the following examples of questions they considered "leading": "Did anything scary happen at naptime?" and "Did anyone ever touch you in a bad place at naptime?"[178] Both Warren and Lamb would consider these questions "specific," and Lamb would characterize them as "leading" but not "suggestive." If, as Bruck and Ceci suggest,[179] interviewers should avoid these types of questions, then the observational research supports the new wave's claims that real-world interviewing is unduly suggestive.

Even in new-wave scholarship, however, what constitutes a "leading" question varies depending on the context. When discussing the real world, the new wave uses the term broadly. On the other hand, in describing their own research, the new wave uses the term quite narrowly. For example, recall the Sam Stone Study. Ceci and Bruck described the study as one in which, "[d]uring each interview, the children were asked two leading questions."[180] Four such interviews, combined with stereotype induction, led to a startling seventy-two percent of three- to four-year-olds making false claims about Sam Stone.

A close examination of the questions reveals that they were much more than merely "leading" or even "suggestive" in Lamb's use of the term. Rather, the questions were "suppositional" or "highly mislead-

---

178 Bruck & Ceci, *supra* note 39, at 306 (internal quotation marks omitted); *see also* Testimony of Maggie Bruck at 15,314, State v. Kelly, No. 933SC676 (N.C. Super. Ct. 1992) ("[I]f I spilled my glass of water and you said, Did you spill your glass of water? That's a leading question because the information is in the question."). Note that although the questions arguably provide some information about the interviewer's hypothesis (that something scary happened at naptime), they fail to name the alleged perpetrator or describe the alleged acts in any detail. More recently, Bruck and her colleagues refer to "leading" questions as those in which the "question stem presupposes the desired answer." Bruck et al., *supra* note 30, at 140. I am unsure how they would classify the aforementioned questions using this definition.

179 *See* Testimony of Maggie Bruck at 15,314, *Kelly*, No. 933SC676 (explaining that leading and misleading questions "are questions that we have been warned not to use with children because it's very easy for the kids to kind of give a response just based on the information that's in the question, not based upon the information that's in their head"); Bruck & Ceci, *supra* note 39, at 306 (acknowledging but not endorsing the views of "some interviewers" who "advocate the use of leading questions as a last resort"); Ceci & Bruck, *supra* note 138, at 18 ("Interviewers who ask nonleading questions, who do not have a confirmatory bias (i.e., an attachment to a single hypothesis), and who do not repeat close-ended, yes/no questions within or across interviews, are more likely to obtain accurate reports from children.").

180 Ceci & Bruck, *supra* note 9, at 416. Later, the authors refer to the questions as "suggestive," which they elsewhere treat as synonymous with "leading." *See, e.g., id.* at 411 ("A second concern is the number of suggestive questions included in the interviews. For example, the Marin et al. (1979) study included only one leading question."); *see also* Michelle D. Leichtman et al., *The Nature and Development of Children's Event Memory, in* TRAUMA AND MEMORY 158, 171 (Paul S. Applebaum et al. eds., 1997) (describing suggestions as "erroneous suggestions" and as "leading questions").

ing."[181] For example, interviewers asked the children questions like "When Sam Stone ripped the book, did he do it because he was angry, or by mistake?"[182] In such a question, the interviewer does not merely introduce the subject of ripping the book, nor does the interviewer merely attempt to obtain the child's acknowledgment that Sam Stone ripped the book. Rather, the question contains the premise that Sam Stone actually ripped the book, and without enabling the child to deny the premise, asks the child for elaboration on the story.

Does this type of questioning make a difference? The new wave acknowledges that it does. Bruck and colleagues cited work from the turn of the century by Lipmann and Wendriner, who "found that preschoolers were progressively susceptible as the strength of misleading questions was increased."[183] The question "Is the door open in the cabinet in the room?" elicited almost ten times as many false affirmations as the question "Is there a cabinet in the room?"[184] Note that the former question resembles those interviewers asked in the Sam Stone Study, whereas one could consider the latter "leading" under the observational research.[185]

Examination of the suggestive interviews in the other studies of the new-wave research reveals similarly strong manipulations. Recall the Inoculation Study, in which researchers falsely "reminded" the children in three interviews that they had not cried when they received their shot one year previously, and that the roles of the research assistant and pediatrician had been switched.[186] What did

---

181 This Article borrows the term "highly misleading" from Ceci and Bruck, who use it to describe questions used in Binet's research that are analogous to those used in the Sam Stone Study (e.g., "What was the color of the thread that attached the button to the board?"). CECI & BRUCK, *supra* note 36, at 54 (internal quotation marks omitted). In the official write-up of the Sam Stone Study, the authors referred to the questions as "erroneous suggestions." Leichtman & Ceci, *supra* note 34, at 571.

182 Leichtman & Ceci, *supra* note 34, at 578 (internal quotation marks omitted).

183 Bruck et al., *supra* note 30, at 139.

184 *Id.* (internal quotation marks omitted).

185 Indeed, the contention that some leading questions are more leading than others is probably one of the oldest and most often replicated findings in the suggestibility literature. *See, e.g.*, Goodman, *supra* note 9, at 19 (discussing Binet's turn-of-the-century work, which found differences in suggestiveness among the following leading questions: "How is the button fastened?"; "Is the button fastened with a thread?"; and "What is the color of the thread which passes through the hole of the button and fixes it to the card?" (internal quotation marks omitted)). Ceci and Bruck have discussed the step-wise effects of increasingly leading questions elsewhere. They cite as leading, but not necessarily deleterious, questions such as "Is he bad because he takes things that do not belong to him, or because he doesn't share things with others, or is he bad for some other reason?" CECI & BRUCK, *supra* note 36, at 296 (internal quotation marks omitted). They find questions like "Is he bad because he touches your private parts?" to be worse, and questions like "He's bad because he does things to your private parts, doesn't he?" to be worst of all. *Id.* (internal quotation marks omitted). Note that the Sam Stone questions go even further because the child is not provided the option of simply saying "no."

186 *See* CECI & BRUCK, *supra* note 36, at 109; Bruck & Ceci, *supra* note 39, at 280.

these "reminders" entail? The interviewer told each child that she and the research assistant worked with the child's pediatrician and that the research assistant was at the office the day that the child received her shot. The interviewer then asked the child to pick out pictures of the pediatrician and the research assistant. In sixty-five percent of the cases, the interviewer had to show the child the correct photograph.[187] Following this exchange, the interviewer kept the photographs in sight during the remainder of the interview and during the interviews that followed.[188] The interviewer then told the child that the research assistant "gives kids their shots. She gave you your shot. Laurie said that she remembered when she gave you your shot . . . ."[189] The interviewer thus asserted the suggested information as a rule, as a specific fact, and as a remembered fact by the alleged actor. The questions that the interviewer asked the children were akin to those in the Sam Stone Study, in which the suggested information served as the question's premise, and the child was asked to provide additional details.[190] Compared to the interviewer's explicit assertions of knowledge in the Inoculation Study, however, the Sam Stone questions seem mild.

In its latest research, the Monkey-Thief Study, the new wave has gone far beyond leading questions in an attempt to suggest nonevents to preschool children. In the suggestive interviews, the interviewer clearly told the children she believed the events had occurred and related information about the event that other children allegedly had provided.[191] If a child denied that the event occurred, the interviewer asked the child to pretend that it had occurred and then asked "specific" questions about the nonevent.[192] In addition, the interviewer employed "visualization techniques, repeating misinformation, and selective reinforcement."[193] Because of the number of suggestive techniques employed, the study produced among the highest percentage of false assents. At the same time, however, the research is likely the least generalizable to the real world.

The Mousetrap Study did not use leading questions.[194] In this study, an examiner asked children questions seven to ten times over a ten-week period. The experimenter simply asked the child to "think

187 *See* Bruck et al., *I Hardly Cried*, *supra* note 34, at 200.

188 *See id.* at 200-01.

189 *Id.* at 200 (internal quotation marks omitted).

190 *See id.* at 201. For example, the interviewer asked the children, "When Laurie (RA) gave you the shot, was your mom or dad with you?" *Id.* (internal quotation marks omitted).

191 *See* Bruck et al., *supra* note 108, at 204.

192 *See id.* The authors do not provide an example of a "specific" question.

193 *Id.*

194 *See* CECI & BRUCK, *supra* note 36, at 218-19.

real hard" about whether each event had occurred, reading the event off a card.[195] The study limited the manipulation to the number of times that the interviews were conducted.[196] Although a large number of children responded incorrectly, the authors did not find an increase in false affirmations over the course of the seven to ten interviews.[197] In a subsequent version of the study—the Bicycle Study—the interviewer told the children that the events had in fact occurred and helped them to imagine relevant details.[198] In the Bicycle Study, false affirmations did increase over the course of the study.[199] Taken together, the two studies present a compelling demonstration of the difference between merely asking a child if an event occurred and telling the child that the event occurred.

In part, telling rather than asking implies that the interviewer knows what occurred, which increases the likelihood that the child will accept the interviewer's suggestions. Ceci and colleagues demonstrated that preschool children are more suggestible when an adult questions them than when a child questions them.[200] They attributed

---

195 *Id.* at 218 (internal quotation marks omitted).

196 *See id.*

197 *See* Ceci et al., *supra* note 103, at 397 ("The results of this study demonstrate that while it is possible to mislead young children into claiming that they experienced nonevents, the frequency of doing so does not increase over time.").

198 *See* CECI & BRUCK, *supra* note 36, at 220.

199 *See* Ceci et al., *supra* note 107, at 311. Unfortunately, readers of Ceci and Bruck's book might overlook the fact that false affirmations did not increase over time in the Mousetrap Study because the authors describe the findings of the Bicycle Study as consistent with the Mousetrap Study: "*As in the previous study,* with each session children increasingly assented to false events." CECI & BRUCK, *supra* note 36, at 221 (emphasis added); *see also* Stephen J. Ceci et al., *Children's Reports of Personal Events, in* Developmental Perspectives on Trauma 515, 519 (Dante Cicchetti & Sheree L. Toth eds., 1997) (discussing the Mousetrap Study and noting that "the mere act of repeatedly imagining participation in an event *caused* these preschoolers to falsely report that they had engaged in the events when they had not" (emphasis added)); Leichtman et al., *supra* note 180, at 173 (discussing the Mousetrap Study and noting that "by the final interview, more than a third of the children reported remembering an event that never occurred, and in most cases these were events that they had denied remembering earlier"). Unfortunately, press reports reinforce possible misconceptions about the Mousetrap Study. *See Morning Edition, supra* note 46 (interviewing Ceci, who reports that "initially when you ask children ages three to six this question, ninety-plus percent get it right," but "[b]y the 10th, 11th week, the majority of 3- and 4-year-olds will claim that getting their hand caught in a mouse trap really happened"); *20/20: From the Mouths of Babes, supra* note 46 ("At first all the kids say no, but then, once a week for 10 weeks, they ask the question again—no coercion, no leading questions as in the child abuse cases. They just gently repeat the question. . . . By week four or six or 10, most of the kids are saying, 'Yes, it happened.'"). In a recent review of their research, Ceci and Huffman appear to have referred inadvertently to a figure depicting data from the Bicycle Study in their discussion of the Mousetrap Study, making it appear that false assents increased over interviews. *See* Ceci & Huffman, *supra* note 34, at 953 fig.7; *cf.* CECI & BRUCK, *supra* note 36, at 221 fig.14.1 (presenting a figure virtually identical to figure 7 referenced to the Bicycle Study).

200 *See* Stephen J. Ceci et al., *Suggestibility of Children's Memory: Psycholegal Implications,* 116 J. EXPERIMENTAL PSYCHOL.: GEN. 38, 42-43 (1987); *see also* Myunghi S. Kwock & Gerald

this difference to the young child's assumption that adults are more credible than other children.[201] Subsequently, Ceci and colleagues obtained similar results with four- to eight-year-olds when both interviewers were adults but varied with respect to how knowledgeable they seemed. One interviewer appeared uncertain of the story about which he suggested information, stating that he did not "remember very much about the story because [he] had not read it in a long time."[202] The other interviewer claimed to know the story "'real well.'"[203] Children were only about half as suggestible when questioned by the interviewer who claimed to know little about the story.[204]

Whether one adopts a narrow or a broad definition of "leading" questions and "suggestive" interviewing also affects the interpretation of other suggestibility research. For example, Saywitz's genital examination study found a three percent to six percent false affirmation rate for anal and vaginal touch among five- to seven-year-old girls.[205] How do Saywitz's questions compare to the observational research on real-world interviews? After first asking for free recall, the researchers asked children seventy questions, most of which were specific questions that called for a yes or no or for a single word.[206] Thirty percent of the questions were suggestive in that the interviewer either presupposed the truth of the question (e.g., "How many times did the doctor kiss you?") or clearly implied a preferred answer (e.g., "She took her clothes off, didn't she?").[207] The results are comparable to the observational research; indeed, the Saywitz study contained a much higher proportion of questions that were suggestive than the Lamb studies.[208] Nevertheless, Ceci and Bruck doubted the extent to which one may generalize from the Saywitz study to the real world.[209] Ceci and Bruck argued that the Saywitz study provided the "optimal conditions under

---

A. Winer, *Overcoming Leading Questions: Effects of Psychosocial Task Variables*, 78 J. EDUC. PSYCHOL. 289, 291-292 (1986) (noting that third graders, but not sixth graders, were better able to overcome the misleading implications of questions from peers than of questions from adults).

201 *See* Ceci et al., *supra* note 200, at 42-43.

202 Michael P. Toglia et al., *The Suggestibility of Children's Memory: A Social-Psychological and Cognitive Interpretation*, *in* DEVELOPMENT OF LONG-TERM RETENTION 217, 225 (Mark L. Howe et al. eds., 1992) (italics and internal quotation marks omitted).

203 *Id.*

204 *See id.* at 224-26 (noting an 18% error rate for children interviewed by the "non-credible" interviewer versus a 33% error rate for children interviewed by the "credible" interviewer).

205 *See* Saywitz et al., *supra* note 59, at 685.

206 *See id.* at 684.

207 *Id.* (internal quotation marks omitted). The authors refer to these as "misleading" questions. *Id.*

208 *See supra* text accompanying note 175 (explaining that approximately 10% of questions in Lamb and colleagues' observational research were suggestive).

209 *See* CECI & BRUCK, *supra* note 36, at 73.

which children should be interviewed."[210] Less charitably, Bruck dismissed the study as "meaningless . . . in terms of what goes on in sexual abuse cases."[211] With no support from the observational research on real-world interviews, Ceci and Bruck raised claims regarding the neutral emotional tone of Saywitz's interviews and the effects of incorporating non-abuse-related questions into the interviews.[212] The effects of a neutral emotional tone or neutral questions on suggestibility, and the extent to which they appear in real-world interviews, are largely unknown.

In sum, "leading" questions are certainly common in investigative interviews. Yet the new wave's research goes far beyond asking leading questions in assessing children's suggestibility. The limited observational evidence available indicates that the new-wave methods are not, in fact, common investigative techniques. The experimental evidence indicates that such techniques are largely responsible for the impressive demonstrations of suggestibility that the new wave has produced.

D. Leading Questions in Court

Suggestibility research rarely examines the effects of countersuggestions—questions that suggest to the child that an event did *not* occur. However, every defendant has the right to ask such questions. While prosecutors must obtain special permission to ask leading questions of children on direct examination, defense attorneys may ask them as a matter of course on cross-examination.[213] Moreover, observational research examining child sexual abuse trials confirms that "the routine use of the leading question [is] very much the preserve of the defense."[214]

When interviewers use countersuggestions in research interviews, children tend to capitulate quickly. For example, the researchers in the Sam Stone Study "gently challenged" children who falsely reported that Sam Stone had performed the suggested misdeeds, and

---

210 *Id.*

211 Testimony of Maggie Bruck at 15,458-59, State v. Kelly, No. 933SC676 (N.C. Super. Ct. 1992).

212 *See* CECI & BRUCK, *supra* note 36, at 73. Saywitz and colleagues discussed the potential limitations of their study in Saywitz et al., *supra* note 59, at 690-91.

213 *See* FED. R. EVID. 611(c) ("Leading questions should not be used on the direct examination of a witness except as may be necessary to develop the witness's testimony. Ordinarily leading questions should be permitted on cross-examination.").

214 GRAHAM DAVIES & ELIZABETH NOON, AN EVALUATION OF THE LIVE LINK FOR CHILD WITNESSES 62 (1991) (examining trials in England and Wales and concluding that 32% of cross-examinations "us[ed] leading questions almost exclusively" as compared with two percent of the examinations in chief).

the rate of false affirmations dropped by half.[215] In another study, Crossman repeatedly and suggestively interviewed three- and four-year-olds in order to convince them that they had witnessed a woman steal money from a purse.[216] Practicing trial attorneys then questioned the children for no more than twenty minutes through direct and cross-examination in a mock courtroom and in the presence of the "defendant." Even under direct examination, children were reluctant to make an allegation,[217] and under cross-examination, only one of the five children who claimed that the defendant stole the money actually testified that he saw her take it.[218]

In the Mousetrap Study, the authors informally debriefed children who had provided false reports. In contrast to the aforementioned results, Ceci and Bruck state that twenty-seven percent "of the children in [the Mousetrap] study refused to accept our debriefing, insisting that they remembered the fictitious events occurring."[219] At first glance, this figure suggests that a fourth of three- to six-year-olds form unshakable memories after repeated interviewing. However, the figure could refer to twenty-seven percent of the children who *consistently affirmed nonevents,* and not to twenty-seven percent of the *entire sample.* As Ceci and his colleagues explained in their original research report, they were only able "to reinterview some of the children who had consistently made false assents," and most of them ultimately accepted that the event had never occurred.[220]

---

215 *See* Leichtman & Ceci, *supra* note 34, at 572, 573 (noting that 44% of three- to four-year-olds claimed they saw Sam Stone perform one or both of suggested acts and 21% did so "even when gently challenged with a countersuggestion"). Unfortunately, descriptions of the study sometimes omit the effects of the countersuggestion. *See* Leichtman et al., *supra* note 180, at 172 ("[A] subset of subjects was also asked as a final question whether they actually saw Sam do the misdeeds they asserted with their own eyes. In response, . . . 44% of the[ ] children in the younger group said that they actually saw him do these things.").

216 *See* Angela M. Crossman, Cross-Examination: Friend or Foe of the Child Witness? (1997) (unpublished manuscript, on file with author) (presented at the Biennial Conference of the Society for Research in Child Development).

217 *See id.* Crossman noted that "children's 'no' responses reached their highest number during direct examination (45.5%), just above the 45% rate of the first interview." *Id.*

218 *See id.*

219 CECI & BRUCK, *supra* note 36, at 220; *see also* Ceci & Huffman, *supra* note 34, at 953 ("Neither parents nor researchers were able to convince 27% of the children that the events never happened.").

220 Ceci et al., *supra* note 103, at 397. A total of nine children were reinterviewed. A reporter from the television program *20/20* conducted one of these interviews, and the researchers conducted the other eight. *See id.* at 400 & n.7. Three of the nine children protested "strongly" to debriefing, three protested "mildly," and three accepted the countersuggestions. *See id.* This research suggests that the figure discussed in the text accompanying note 219 *supra* should be 33% and not 27%. I have been unable to reconstruct the exact origin of the 27% figure.

In addition to asking leading questions, defense attorneys have other means by which they can undermine a child witness's credibility on the stand. Two-thirds of defense attorneys in one survey admitted that they would "often" or "always" "use to advantage the child's vulnerabilities during cross-examination."[221] Although several commentators have noted that brutalizing a child on the stand is poor strategy because doing so makes the child more sympathetic in the jury's eyes,[222] an attorney need not be brutal to cross-examine effectively a young child. As Crossman found in her mock trial study, "[w]hereas the volunteering attorneys in this case were good communicators experienced with children, they were also good lawyers adept at gently highlighting the inconsistencies and weaknesses in the children's stories, discrediting them with alarming ease."[223]

Reviewing the observational research on attorneys' actual behavior in the courtroom, Jean Montoya argued that "[i]t follows from this research that the assertions in the literature that the defense seeks to intimidate the child witness into silence are uninformed."[224] Montoya pointed to two studies that failed to find differences between child witnesses' demeanor (involving subject ratings of happiness, competence, and credibility) during direct examination and cross-examination.[225]

Examining children's demeanor, however, does not tell the whole story. Montoya acknowledged one study that found cross-examiners less supportive of child witnesses.[226] Indeed, four other studies

---

221 Michael R. Leippe et al., *The Opinions and Practices of Criminal Attorneys Regarding Child Eyewitnesses: A Survey, in* PERSPECTIVES ON CHILDREN'S TESTIMONY 100, 117 (S.J. Ceci et al. eds., 1989) (internal quotation marks omitted) (describing the results of a survey of 74 defense attorneys in Florida).

222 *See* ELLEN GRAY, UNEQUAL JUSTICE: THE PROSECUTION OF CHILD SEXUAL ABUSE 154-55 (1993) ("There is some evidence that defense tactics have softened somewhat for fear that their confrontational style would engender sympathy for the child and hurt the defendant's case."); MCGOUGH, *supra* note 13, at 291 n.14 ("Many seasoned defense attorneys also advise that brutalizing a child witness, even one suspected of lying, is counterproductive.").

223 Crossman, *supra* note 216.

224 Jean Montoya, *Lessons from* Akiki *and* Michaels *on Shielding Child Witnesses*, 1 PSYCHOL. PUB. POL'Y & L. 340, 351 n.78 (1995).

225 *See id.* (citing GAIL S. GOODMAN ET AL., TESTIFYING IN CRIMINAL COURT 87 (Monographs of the Soc'y for Research in Child Dev. No. 229, 1992)); Rhona Flin et al., *Children in the Witness Box, in* CHILDREN AS WITNESSES 176 (Helen Dent & Rhona Flin eds., 1992) [hereinafter Flin et al., *Witness Box*]; Rhona Flin et al., *Child Witnesses in Scottish Criminal Trials*, 2 INT'L REV. VICTIMOLOGY 309, 326 (1993) [hereinafter Flin et al., *Scottish*] (describing the same research as Flin et al., *Witness Box, supra*)); *see also* GRAY, *supra* note 222, at 151-55 (describing similarity of prosecution and defense styles of questioning child witnesses).

226 *See* Montoya, *supra* note 224, at 351 n.78 (citing GOODMAN ET AL., *supra* note 225, at 79-80).

have come to the same conclusion.[227] Studies also have found, albeit less consistently, that cross-examiners tend to ask questions that are more difficult,[228] less age-appropriate,[229] and less accommodating to the child's linguistic style.[230]

One might respond that differences in cross-examination styles are irrelevant if children's performances are unaffected. Yet when differences in performance do appear, children inevitably perform worse on cross-examination than on direct.[231] Furthermore, one must doubt whether the research provides a sufficiently sensitive test of the effects of cross-examination upon young children. Recall Crossman's finding that examiners easily and gently manipulated three- and four-year-old children into contradiction and denial.[232] The observational research involves children whose average age is closer to eleven years of age,[233] with only a small proportion of children under nine years of age.[234] The observational research thus may understate the negative effects of cross-examination on young children.

Another problem is that global and subjective impressions of children's demeanor might fail to reveal real psychological differences,

---

227 *See* JUDY CASHMORE, THE USE OF CLOSED-CIRCUIT TELEVISION FOR CHILD WITNESSES IN THE ACT 59 (1992); DAVIES & NOON, *supra* note 214, at 61; GRAY, *supra* note 222, at 158-59; Flin et al., *Scottish*, *supra* note 225, at 325. Montoya cited all but Gray's study at other points in the paper.

228 *See* CASHMORE, *supra* note 227, at 48-53 (finding that defense asked more difficult questions, and defense questions were less well understood by child wituesses); Flin et al., *Scottish*, *supra* note 225, at 322-23 (finding that "12% of examinations-in-chief, and 40% of cross-examinations contained some vocabulary that the child appeared not to understand").

229 *See* GRAY, *supra* note 222, at 153, 158 (noting the same result at competency evaluations but not at trials); GOODMAN ET AL., *supra* note 225, at 80 (finding defense questions less age-appropriate at both preliminary hearings and trials).

230 *See* Davies & Noon, *supra* note 214, at 58.

231 *See* CASHMORE, *supra* note 227, at 48-53 (finding that children misunderstand more when questioned by defense); DAVIES & NOON, *supra* note 214, at 59 (finding more "I don't know's" and more inconsistencies when questioned by the defense); GOODMAN ET AL., *supra* note 225, at 81 (finding that children answered fewer questions at preliminary hearing and at trial when questioned by the defense); Flin et al., *Scottish*, *supra* note 225, at 326 (finding that children are less confident when questioned by the defense). Limited evidence also suggests that cross-examination has a greater effect on younger children than older children. *See* GOODMAN ET AL., *supra* note 225, at 82, 91 (finding that younger children faltered more than older children on cross-examination at preliminary hearing but not at trial, although noting that the sample contained only 17 cases that went to trial).

232 *See supra* text accompanying notes 216-18.

233 *See* CASHMORE, *supra* note 227, at 13 (reporting a mean age of 11.8 with a range of five to 17); DAVIES & NOON, *supra* note 214, at 22 (reporting a mean age of 10); GOODMAN ET AL., *supra* note 225, at 78, 87 (noting that the mean age at preliminary interviews was nine years and nine months with an age range of four to 15 and the mean age at trial was 11 years, four months with a range of five to 17). GRAY, *supra* note 222, and Flin et al., *Scottish*, *supra* note 225, did not report mean ages.

234 *See* DAVIES & NOON, *supra* note 214, at 22 (noting only eight percent under seven years of age); Flin et al., *Witness Box*, *supra* note 225, at 170 (noting only 11% of the children five to eight years of age).

particularly when the trial experience itself is extremely stressful for children regardless of who is conducting the examination.[235] A study examining the differences between children's demeanor under direct examination and cross-examination through closed-circuit television, which is less stressful than testimony in open court,[236] provides indirect support for this proposition. As Montoya acknowledged, the children in that study appeared more unhappy, less effective, and less credible on cross-examination than on direct.[237] An overall reduction in stress may have allowed relatively insensitive measures of demeanor to detect real differences between the effect of direct and cross-examination on children's demeanor.

Partly because of cross-examination, children's performances in court are likely to be quite unlike their performances in the lab. Defense attorneys may use leading questions—countersuggestions—to undermine children's confidence and reduce their credibility in the eyes of the jury. The available experimental evidence indicates that if children are suggestible, they are also countersuggestible. In other words, if leading questions can create false allegations, they also can undermine them.

IV

THE REAL WORLD OF SEXUAL ABUSE: MOTIVATIONAL DISINCENTIVES TO CLAIMING ABUSE

One must understand children's feelings about and reactions to sexual abuse to make policy recommendations regarding the appropriate interviewing of suspected abuse victims and to assess the credibility of children's abuse allegations. If these feelings and reactions motivate children to deny that abuse occurred, at least two implications follow. First, interviewers must move beyond open-ended questioning to overcome abused children's fear and embarrassment. If interviewers avoid leading questions at all costs, one cost will be abused children who withhold details of their abuse. Second, researchers and policymakers debating the ecological validity of suggestibility research must recognize that false allegations of sexual abuse

---

235 *Cf.* Flin et al., *Witness Box*, *supra* note 225, at 177 (arguing that a generalized fear of testifying might mask the differences in children's demeanor between direct and cross).

236 *See* DAVIES & NOON, *supra* note 214, at 72 (finding that children testifying via closed-circuit television "were less unhappy" while testifying than children testifying in open court).

237 *See* Montoya, *supra* note 224, at 351 n.78; *see also* Davies & Noon, *supra* note 214, at 51 (indicating that children were more unhappy when questioned by defense); *id.* at 60 (indicating that children were less effective when questioned by defense); *id.* at 61 (indicating that children are less credible when questioned by defense). However, children did not look uniformly worse on cross-examination because some comparisons did not find statistically significant differences. *See id.* at 52-53 (noting that children were no more tense nor less confident nor less fluent when questioned by defense).

may be more difficult to elicit than the false allegations the new wave produces.

Surprisingly, the new wave calls into question the intuitive claim that children are reluctant to disclose abuse. Bruck and her colleagues recently challenged the "stubborn urban legend" that "when directly asked about abuse, it is common for sexually abused children to not readily or consistently disclose their abuse."[238] The authors pointed out that research documenting a relatively high rate of initial denial (seventy-five percent) is suspect because "the children in these studies may not have been sexually abused."[239] They cited research on substantiated sexual abuse cases that found a much lower rate of denial (five percent) and concluded that "although a small percentage of youngsters do appear to disclose their abuse reluctantly, . . . the overwhelming majority of children appear to maintain their claims and never deny them to officials once they are questioned."[240]

Bruck, Ceci, and Hembrooke overlooked a simple fact about substantiated sexual abuse cases: without a child's statement that abuse occurred, abuse is unlikely to be substantiated.[241] Therefore, studies of substantiated cases necessarily will underestimate abused children's reluctance to disclose. Moreover, the authors ignored research that overcomes this basic methodological flaw and that previously had led them to acknowledge that "truly abused children are often unlikely to disclose sexual abuse out of a sense of embarrassment or fear."[242] Such research examines cases in which evidence other than the child's statements established that sexual abuse occurred. Lawson and Chaffin found that fifty-seven percent of children with a sexually transmitted disease failed to disclose abuse when questioned.[243] Muram and his colleagues found that forty-nine percent of children with medical evidence strongly indicative of sexual abuse failed to dis-

---

238 Bruck et al., *supra* note 30, at 138.

239 *Id.*

240 *Id.* The authors also criticize claims that children often recant their abuse, a position I respond to elsewhere. *See* Thomas D. Lyon, *Scientific Support for Expert Testimony on Child Sexual Abuse Accommodation Syndrome, in* THE KNOWNS AND UNKNOWNS OF CHILD SEXUAL ABUSE (Jon Conte ed., forthcoming 1999).

241 *See* Mary E. Haskett et al., *Substantiation of Sexual Abuse Allegations: Factors Involved in the Decision-Making Process*, 4 J. CHILD SEXUAL ABUSE 19, 40 (1995) (noting that a survey of social workers found that "[b]y far, the most important factor in [the substantiation] process was the child's verbal disclosure or denial of abuse").

242 Ceci et al., *supra* note 33, at 506 (describing this fact as "a point of no dispute among researchers"). The authors argue that such research does not prove that nonabused children are "impervious to the types of suggestive interviewing tactics used by some of the *Michaels* investigators." *Id.*

243 *See* Louanne Lawson & Mark Chaffin, *False Negatives in Sexual Abuse Disclosure Interviews: Incidence and Influence of Caretaker's Belief in Abuse in Cases of Accidental Abuse Discovery by Diagnosis of STD*, 7 J. INTERPERSONAL VIOLENCE 532, 537 (1992).

close abuse.[244] In addition to these studies, which Ceci and his colleagues have acknowledged,[245] at least three other studies found high rates of nondisclosure among children for whom strong external evidence of abuse existed.[246]

Of course, one cannot easily determine the reasons why abused children fail to reveal their abuse, in part because one cannot ask silent victims the reasons for their silence. If a child belatedly reveals her abuse, however, one can ask why she delayed. Similarly, one can ask adults who acknowledge childhood abuse for the first time why they kept their secrets for so long. As we shall see, the following are at least three reasons for this silence: fear, loyalty, and embarrassment. Children may fail to report abuse due to their fear of the potential negative consequences to themselves and their loved ones (including, in some cases, the perpetrator). If the offender is someone close to the child or is a member of the child's family, loyalty to the offender makes the child particularly reluctant to report her abuse. Finally, abused children worry that others will blame them for the abuse, causing feelings of both embarrassment and shame.

Emotions like fear, loyalty, and embarrassment are largely absent from the laboratory research documenting high rates of suggestibility, in part because ethical considerations limit what researchers are allowed to inflict upon their subjects.[247] The absence of the strong emotions that typically accompany abuse allegations limits the ecological validity of suggestibility research. Moreover, the presence of powerful disincentives to disclosure in actual abuse cases may explain why

---

244 *See* David Muram et al., *Genital Abnormalities in Female Siblings and Friends of Child Victims of Sexual Abuse*, 15 CHILD ABUSE & NEGLECT 105, 108 tbl.2 (1991).

245 *See supra* note 242.

246 In a number of studies, the sample can be broken down into groups of children for whom the external evidence of sexual abuse was strong, obviating the bias created by only examining cases that were substantiated by the interview itself. *See* Howard Dubowitz et al., *The Diagnosis of Child Sexual Abuse*, 146 AM. J. DISEASES CHILDREN 688, 691 (1992) ("Since 25% of [the 28] children with abnormal examination findings indicative of abuse did not disclose at all and 28% partially disclosed, without the examination many of these children might not have been diagnosed as abused."); Diana M. Elliot & John Briere, *Forensic Sexual Abuse Evaluations of Older Children: Disclosures and Symptomatology*, 12 BEHAV. SCI. & L. 261, 263-65 (1994) (reporting that 39 of 118 or 33% of children for whom external evidence of abuse existed denied abuse when evaluated at a sexual abuse crisis center, in which external evidence was defined as abnormal medical examinations considered diagnostic of abuse, perpetrator confessions, eyewitness statements, or corroborative evidence, such as pornographic pictures of the child); Stacy Gordon & Paula K. Jaudes, *Sexual Abuse Evaluations in the Emergency Department: Is the History Reliable?*, 20 CHILD ABUSE & NEGLECT 315, 319 (1996) (noting that 24% of children with a sexually transmitted disease denied abuse when questioned both in a hospital emergency room and by a sexual abuse interdisciplinary team).

247 *See* CECI & BRUCK, *supra* note 36, at 68 (noting that the challenge for researchers "has been to incorporate questions that ask whether or not sexual actions occurred . . . but to do so in an ethically permissible manner").

real-world investigators feel compelled to move beyond open-ended questions when asking young children about abuse.

A. Fear and Loyalty

Ceci and Bruck argue that experts in child abuse cases should not testify that threats deter abused children from disclosing their abuse because no empirical basis for this "professional 'lore'" exists.[248] On the other hand, Bruck and Ceci readily assert that threats (and bribes) may induce *nonabused* children to claim abuse based on "everything we know about the principles of child development and about principles of punishment and reward."[249] This inconsistency is particularly perplexing given the severity of threats in true and false reports of abuse. Threats *not* to reveal range from "pleas that the abuser would get into trouble if the child told . . . to threats that the child would be blamed for the abuse . . . to ominous warnings that the defendant would hurt or kill the child (or someone he or she loved) if they revealed the abuse."[250] Ceci and Bruck consider it inappropriately threatening to tell a nondisclosing child: "Don't be a baby. You're acting like a nursery school kid."[251]

1. *Fear in the Lab*

Ceci and Bruck base their reluctance to believe that threats may deter disclosure in part on laboratory research examining children's willingness to keep an adult's transgressions a secret.[252] They highlight the work of Doug Peters, who exposed four- to ten-year-old children to a stranger who stole a book in the presence of each child and then asked the child to keep the theft a secret.[253] Ceci and Bruck summarize the study as demonstrating that "although children in a laboratory experiment would not disclose a crime to their parents if

---

248 *Id.* at 300.

249 Bruck & Ceci, *supra* note 39, at 282.

250 SMITH & ELSTEIN, *supra* note 15, at 93; *see also* HERMAN, *supra* note 10, at 88 (indicating that many victims of incest "were threatened with the most dreadful consequences if they told: their mothers would have a nervous breakdown, their parents would divorce, their fathers would be put in jail, or they themselves would be punished and sent away from home").

251 CECI & BRUCK, *supra* note 36, at 145 (internal quotation marks omitted).

252 *See id.* at 145 n.1.

253 *See* Douglas P. Peters, Confrontational Stress and Children's Testimony: Some Experimental Findings (Mar. 16, 1990) (unpublished manuscript); Douglas P. Peters, Confrontational Stress and Children's Testimony: Some Experimental Findings (Apr. 21, 1991) (unpublished manuscript), *cited in* McGOUGH, *supra* note 13, at 91. Because these studies are unpublished (and I have not been able to obtain copies), I rely on the description of the research in McGOUGH, *supra* note 13. Although there are two papers, McGough implies they are the same study. *See* McGOUGH, *supra* note 13, at 91, 289 n.13.

the perpetrator was present, they were quite likely to do so as soon as the perpetrator was absent."[254]

Ceci and Bruck's conclusion that children are "quite likely" to disclose is based on the fact that after both the owner of the book and their parents questioned the children, sixty-seven percent of the children ultimately disclosed the thief's identity.[255] Others have found it remarkable that even after such questioning, "nearly one-third (32.5 percent) of the children still feigned ignorance."[256] Indeed, another reading of Peters's research highlights the child's strong reluctance to implicate an adult in wrongdoing, even when the adult is a stranger who does not threaten or bribe the child, and when the questioners are sure both that a crime had occurred and that the child had witnessed it. Consider this description of Peters's research:

> Four- to ten-year-olds witnessed a staged event of a stranger who stole a book and were asked to keep the theft a secret. When the children were asked by the owner of the book whether they had seen who took it, 82% either delayed reporting the theft or never reported it. The most common reason given by the children for not disclosing was to honor the stranger's secret and to avoid getting him into trouble.[257]

Ceci and Bruck described Peters's research in this manner in 1993, justifying their contention that "even very young children sometimes do lie."[258]

Although Ceci and Bruck omit the discussion of Peters's work as documenting a reluctance to disclose transgressions from their 1995 book, they discuss a number of studies that provide "consistent evidence that children as young as 3 years of age will omit important information about transgressions and accidents if adults ask them to do so,"[259] and they acknowledge that this evidence "could also be used to address the issue of the degree to which children withhold the truth when they are threatened."[260] In addition to Peters, two groups of researchers have conducted four studies which demonstrate that a substantial percentage of three- to ten-year-old children will keep a stranger's transgressions secret when they believe that revealing the

---

254 CECI & BRUCK, *supra* note 36, at 301.

255 *Id.* at 145 n.1.

256 MCGOUGH, *supra* note 13, at 91.

257 Ceci & Bruck, *supra* note 9, at 426. The strength of the secrecy manipulation is unclear. In their 1993 description of Peters's research, Ceci and Bruck stated that the thief "asked" the child not to tell. *Id.; see also* MCGOUGH, *supra* note 13, at 91 (describing the thief's actions in similar terms). In their 1995 book, however, Ceci and Bruck indicated that the thief "told the child[ren]" not to tell. CECI & BRUCK, *supra* note 36, at 145 n.1.

258 Ceci & Bruck, *supra* note 9, at 425.

259 CECI & BRUCK, *supra* note 36, at 263.

260 *Id.* at 264 n.2.

transgression will get the stranger into trouble and the stranger has asked them not to tell.[261]

Experimental evidence supports the common sense claim that threats reduce the willingness of children to disclose. Furthermore, this research suggests that children's reluctance to disclose increases as the intensity of the warning increases.[262] For ethical reasons, psychologists do not threaten children with serious harm in any of their research, but it is reasonable to posit that such threats would be even more effective.[263]

---

261 In a study by Wilson and Pipe involving five-year-olds, a magician performed a number of tricks for the child and then accidentally spilled ink on "magic gloves" that the child was wearing. J. Clare Wilson & Margaret-Ellen Pipe, *The Effects of Cues on Young Children's Recall of Real Events*, 18 N.Z. J. PSYCHOL. 65, 66 (1989). The magician hid the gloves, "saying if they were discovered she (the magician) would be reprimanded and that therefore they should not tell anyone about the inkspill." *Id.* at 66-67. An interviewer then questioned the child, first ten days after the event and then two months after. Initially, the interviewer asked the child to relate everything that the magician did. Ultimately, the interviewer asked the child whether the child knew anything about a pair of stained gloves the interviewer had found. *See id.* at 67. None of the children spontaneously mentioned the gloves after 10 days, and 75% failed to do so after two months. *See id.* at 68. Twenty-five percent denied knowing anything about the gloves at both interviews when directly asked, and another 33% denied knowing anything at one of the two interviews. *See id.* Pipe and Wilson subsequently found similar rates of nondisclosure among six-year-olds and less reluctance to disclose among ten-year-olds. *See* Margaret-Ellen Pipe & J. Clare Wilson, *Cues and Secrets: Influences on Children's Event Reports*, 30 DEVELOPMENTAL PSYCHOL. 515, 518-19 (1994). Most six-year-olds failed to mention the gloves in their free recall (75% at two weeks, 81% at two months), and over 30% failed to reveal what happened after the specific question was asked (40% at two weeks, 32% at two months). *See id.* at 521 tbl.3. The 10-year-olds were less inclined to keep the incident a secret, but nevertheless over 30% failed to mention the gloves in free recall (34% at two weeks, 44% at two months), and 16% did not reveal when specifically asked (at both interviews). *See id.* Bussey and colleagues examined the willinguess of three- and five-year-olds to remain silent about a male experimenter who had accidentally broken a prized glass and hidden the pieces. *See* Kay Bussey, Factors Infiuencing Children's Disclosure of Witnessed Events (Mar. 1993) (unpublished manuscript, on file with author). "He expressed a great deal of concern about the event and sought to dissuade the child from disclosing what had happened as he hid the broken glass under some paper in the bin." *Id.* The female experimenter subsequently asked the child questions about the glass, including "Did [the male experimenter] touch the glass?" if the child had not already revealed this information. *Id.* (internal quotation marks omitted). Fourteen percent of three-year-olds and 43% of five-year-olds kept the secret. If the experimenter sternly told the child not to tell, 43% of the three-year-olds and 71% of the five-year-olds either denied that the mishap occurred or refused to discuss it. *See* Kay Bussey & Elizabeth J. Grimbeek, *Disclosure Processes: Issues for Child Sexual Abuse Victims*, *in* DISCLOSURE PROCESSES IN CHILDREN AND ADOLESCENTS 166, 182 (Ken J. Rotenberg ed., 1995). Bussey reported lower rates of nondisclosure among nine-year-olds (approximately 15% after being asked not to tell). *See* Bussey, *supra*, at 10.

262 *See* Bussey & Grimbeek, *supra* note 261.

263 *See* McGOUGH, *supra* note 13, at 91 ("When the threat to the child is made more explicit, we should expect an even stronger inclination of the child to lie."); Pipe & Wilson, *supra* note 261, at 523. Pipe and Wilson suggested the following:

> Insofar as the events in the studies to date have been relatively neutral, it is likely that they underestimate the willingness of children to omit information when requested to do so in real-life contexts. We can only speculate on how much more likely children will be to maintain secrecy when, for

2. *Parents Versus Strangers*

Ceci and Bruck have argued that "[i]f children will lie to protect a stranger, they should do so even more readily to protect a loved one."[264] Abusers, often family members or friends of the family, are typically close to the children they abuse.[265] A child will have greater sympathy for one she loves and probably is less inclined to expose such a person. If the loved one is in the child's home, or close to others that the child loves, threats and inducements may be even more effective. In these situations, the offender has continuing contact with the child and others in the family, and the child cannot count on being supported by other loved ones should she disclose. Adults who were abused as children mention these concerns when explaining why they never revealed their abuse.[266]

Laboratory research further supports the reluctance of children to implicate parents and others close to them. Ceci and Leichtman have shown that if a researcher spends twenty hours with a three- to four-year-old child, thereby becoming a "loved one," that child will strongly resist revealing wrongdoing by the researcher.[267] In an analo-

---

> example, they feel embarrassment or guilt, have been asked to conceal information by a person to whom they feel some strong obligation, have given a commitment to secrecy as in a promise, or are faced with threats to themselves or their family, if they disclose.

Id.

264 CECI & BRUCK, *supra* note 36, at 264.

265 *See supra* note 143 and accompanying text.

266 *See* HERMAN, *supra* note 10, at 88 ("[T]he girls were given to understand that breaking secrecy would lead to separation from one or both of their parents. Those who remembered no warnings simply intuited that guarding the incest secret was part of their obligation to keep the family together . . . ."); DIANA E.H. RUSSELL, THE SECRET TRAUMA: INCEST IN THE LIVES OF GIRLS AND WOMEN 132 (1986); Robert L. Johnson & Diane K. Shrier, *Sexual Victimization of Boys: Experience at an Adolescent Medicine Clinic*, 6 J. ADOLESCENT HEALTH CARE 372, 374 (1985) (noting that boys "who had not previously revealed the assault" mentioned the desire "to protect the assailant or were afraid of the reactions of their parents or family members"). Diana Russell spelled out these concerns particularly well:

> In those cases where the victim did not tell anyone, we tried to ascertain the primary reason for secrecy. For these forty-four incest victims, the two most common reasons were fear of punishment by the perpetrator and/or someone else, including abandonment or rejection and a desire to protect the perpetrator, or fear of hurting someone else. For other victims self-blame made them feel too ashamed or guilty to tell. Some expressed fear of being blamed or of not being believed.

RUSSELL, *supra*, at 132.

267 Stephen J. Ceci & Michelle DeSimone Leichtman, *"I Know That You Know That I Know That You Broke the Toy": A Brief Report of Recursive Awareness Among 3-Year-Olds*, *in* COGNITIVE AND SOCIAL FACTORS IN EARLY DECEPTION 1, 6 (Stephen J. Ceci et al. eds., 1992). The experimenter and the child were told by a nursery school teacher not to play with a toy. While the teacher was gone, the loved one touched and broke the toy and exclaimed, "'Gee, I didn't mean to break it. I hope I do not get into trouble for breaking this.'" *Id.* Note that the loved one did not elicit a promise from the child nor threaten the child not to tell. The teacher returned and asked the child who broke the toy. "Most children, when

gous study, Devitt and her colleagues compared children's willingness to implicate a parent with their willingness to implicate a stranger. A stranger stole a book in the presence of a child (four to eleven years of age) and told the child "that the theft was to be their secret and that the child should not tell anyone that the researcher had taken the book."[268] The owner of the book discovered it was missing and explained to the child that it was needed for an exam the next day. The owner and an experimenter questioned the child, and the experimenter then asked the child and her parent to wait for the police to arrive. A person identified as an officer then questioned the child. Nineteen percent of the children failed to name the thief.[269] In an alternate condition in which the child watched as his or her *parent* stole the book, and the parent told the child to name one of the experimenters as the thief, eighty-one percent of the children failed to name the thief (fifty-six percent falsely accused the experimenter named by the parent, and twenty-five percent failed to name anyone).[270]

The Devitt study suggests that a parent may be able to create a false allegation. If parents can commit a crime and coach their children to accuse another, perhaps parents in custody battles can coach their children to accuse falsely the other parent of abuse. In order to accept the analogy, however, one must assume that the accused parent is no more than a virtual stranger to the child. In reality, noncustodial parents accused of sexually abusing their children often have

---

confronted with the choice of disclosing that their loved one broke it, either refused to say anything or provided misleading information (e.g., 'A gremlin came in through the window and broke it.')." *Id.*

268 Mary K. Devitt et al., A Study of the Willingness of Children To Make False Accusations About a Serious Matter in a Realistic Setting, Paper Presented at the Biennial Meeting of the American Psychology and Law Society (Mar. 1994), *cited in* Charles Robert Honts, *Assessing Children's Credibility: Scientific and Legal Issues in 1994*, 70 N.D. L. REV. 879, 883 n.20 (1994).

269 *See* Honts, *supra* note 268, at 884.

270 *See id.* at 883 n.20, 884-85. In a study by Bottoms and colleagues involving three- to four-year-olds and five- to six-year-olds, participants were divided into two groups. Bette L. Bottoms et al., Keeping Secrets: Implications for Children's Testimony (Mar. 1990) (unpublished manuscript, on file with author); *see also* Margaret-Ellen Pipe & Gail S. Goodman, *Elements of Secrecy: Implications for Children's Testimony*, 9 BEHAV. SCI. & L. 33, 37 (1991) (discussing Bottoms et al., *supra*). Both groups of children saw their mother accidentally break the head off a Barbie doll. In the secrecy group, the mother and child had been told not to play with the toys, and the mothers "instructed their children to keep the fact they had played with the toys a secret[, telling] her child that she might get in trouble if the child told, and that the child would get to keep one of the toys if he/she kept the secret." Bottoms et al., *supra*, at 5-6. In the control group, the mother and child were free to play with the toys, and the mothers did not give their children any instructions about secrecy. *See id.* Only one of the 49 children in both age groups told an interviewer about the doll when asked what happened, and "when asked specific questions about the event, 5-year-olds did not tell the secret, even when asked leading questions." CECI & BRUCK, *supra* note 36, at 264.

had frequent and extended contact with their children. Indeed, the existence of such contact often forms the basis for the custodial parents' suspicions.

Similarly, Leichtman and Ceci have argued that stereotype induction (the phenomenon studied in the Sam Stone Study)[271] may occur in the context of divorce cases: "[T]he defendant might be an estranged parent who has been previously criticized by the custodial parent in the child's presence, and the child may even have come to accept these criticisms as stable aspects of the parent's character."[272] However, Sam Stone was as much of a stranger to the children as the stranger in the Devitt study. Children were exposed to Sam Stone for two minutes,[273] during which "nothing happened."[274] Sam Stone did not interact with the children individually, but only spoke to the group while they were listening to a story read by a teacher.[275]

In contrast, children became much more familiar with the adults who sought to influence their perception of Sam Stone. Children in the stereotype induction condition played with research assistants for four consecutive weeks before Sam Stone's visit and "received considerable information" about Sam, including descriptions of twelve accidents witnessed by the research assistant that each depicted Sam as an accident-prone person.[276] In the suggestion condition, children individually played with research assistants for four consecutive weeks after Sam Stone's visit.[277] Therefore, children in the stereotype and suggestion condition interacted with research assistants on eight different occasions, whereas they only saw Sam Stone on one occasion.[278]

In addition to the lack of familiarity with Sam, the children had few other incentives to protect Sam Stone. The research assistants did not lead the children to believe that revealing Sam's misdeeds would get him in trouble. On the contrary, they depicted Sam "as a kind, well-meaning, but very clumsy and bumbling person," whose misdeeds were accidental and promptly corrected.[279] Moreover, researchers gave the children little opportunity to deny wrongdoing because during the first two interviews the research assistants presented the children with a ripped book and a soiled teddy bear—remnants of Sam's alleged misdeeds.[280] Presenting a child with evidence of the mishap is

271 *See* Leichtman & Ceci, *supra* note 34, at 570.
272 *Id.* at 569.
273 *See id.* at 570.
274 *Id.* at 570, 575.
275 *See id.* at 570.
276 *Id.*
277 *See id.* at 571.
278 *See id.*
279 *Id.* at 577.
280 *See id.*

implicitly accusatory and motivates the child to name an offender. In contrast, an abused child may deny that anything wrongful has occurred in order to foreclose discovery of something shameful.

If Ceci and Bruck are correct that children are more willing to lie for loved ones, then they also may be more willing to tell the truth for loved ones. If this assertion is true, then it follows that children may resist suggestions better when the accused is a parent. Ceci and his colleagues previously have emphasized the difficulty in extrapolating from studies in which children were "presented short vignettes in an affectively neutral context by unfamiliar adults" to situations in which children testify about events that were "of a repetitive nature (e.g., sexual molestation), in an emotionally charged context, and perpetrated by a familiar person, often a family member."[281] In defending the Sam Stone Study against such criticism, however, Ceci and his colleagues argued that it is a "misrepresentation of the literature"[282] to claim that the study "minimizes the likelihood that children's familiarity with and respect for the alleged wrongdoer militates against their suggestibility."[283]

The authors justified their reluctance to acknowledge the effects of familial ties on suggestibility by referring to a study by Lepore and Sesco[284] and to the infamous mass-abuse cases in which children accused their parents of abuse, such as the Jordan, Minnesota case.[285] In the Lepore and Sesco study, in which many preschool children falsely claimed that a man had taken their clothes off and kissed them, the researchers defined familiarity as one and a half hours of exposure to an individual about whom the researchers asked misleading questions.[286] Compare this exposure to the twenty-one hours Ceci and Leichtman felt necessary to create a "loved one" for their study on

---

281 Ceci et al., *supra* note 200, at 47. Notably, in discussing the "new line of research that has attempted to address these criticisms [including that of Ceci, Ross, and Toglia]," Ceci and Bruck rephrase the question as whether one can mislead children "about very important or salient events, especially those involving their own bodies." CECI & BRUCK, *supra* note 36, at 67. Reference to the relationship between the child and the person with whom the child interacts is omitted.

282 Ceci et al., *supra* note 33, at 506.

283 Thomas D. Lyon, *False Allegations and False Denials in Child Sexual Abuse*, 1 PSYCHOL. PUB. POL'Y & L. 429, 431 (1995).

284 *See* Stephen J. Lepore & Barbara Sesco, *Distorting Children's Reports and Interpretations of Events Through Suggestion*, 79 J. APPLIED PSYCHOL. 108 (1994).

285 *See* Ceci et al., *supra* note 33, at 506.

286 *See* Lepore & Sesco, *supra* note 284, at 109 ("[C]hildren in the familiar-TA condition had seen the TA in their classroom on three occasions in the 2 weeks before this staged interaction."). The authors themselves emphasized the weakness of the familiarity manipulation. *See id.* at 118. Furthermore, although they found that 90 minutes of interaction did not reduce children's susceptibility to an interview containing stereotyping, peer pressure, and suggestive questions, familiarity improved the accuracy of children's responses to open-ended questions regarding their interaction with the stranger and led the children to rate the stranger as more likable. *See id.* at 111 tbl.1, 113.

children's denial of the wrongdoing of others. In the *Jordan* case, one of the relatively few mass-allegation cases in which parents were the accused, interviewers and others told children that they would never see their parents again if they failed to reveal, and that they could help their parents by accusing them of abuse.[287] This case suggests that loyalty engenders strong motives to protect, and that unless one can trick children into believing that accusing their parents of molestation will help their parents, they are strongly motivated not to do so.

Undeniably, a vindictive parent may turn her child against the other parent. However, although nonoffending mothers are more likely to believe their children's allegations when the accused is the ex-husband,[288] a substantial proportion of nonoffending mothers overall are ambivalent or unsupportive of their children's claims.[289]

---

287 *See* CECI & BRUCK, *supra* note 36, at 237 ("It seems that in some cases, children were engaged in highly pressurized interviews; they were told that they could help their (imprisoned) parents after they had been removed from their families or their homes.").

288 *See* DIVISION OF CHILD PSYCHIATRY, TUFTS NEW ENGLAND MED. CTR., SEXUALLY EXPLOITED CHILDREN 196-97 (1984) [hereinafter TUFTS] ("Mothers were least protective and most angry and punitive toward *the child* when the abuser was not the natural father, but a stepfather or boyfriend." (emphasis added)); Allan R. De Jong, *Maternal Responses to the Sexual Abuse of Their Children*, 18 PEDIATRICS 14, 18 (1988) ("Forty-four percent (11 of 25) of abuse by fathers or by the mother's paramours resulted in nonsupportive maternal responses, whereas only 27% (21 of 78) of the assaults by all other perpetrators resulted in nonsupportive responses . . . ."); Mark D. Everson et al., *Maternal Support Following Disclosure of Incest*, 59 AM. J. ORTHOPSYCHIATRY 197, 200 (1989) ("[M]others were significantly more supportive of their children if the offender were an ex-spouse than if he were someone with whom the women had a current relationship."); Elizabeth A. Sirles & Pamela J. Franke, *Factors Influencing Mothers' Reactions to Intrafamily Sexual Abuse*, 13 CHILD ABUSE & NEGLECT 131, 134 (1989) (reporting that "[i]f the offender was a biological father, 85.9% of the mothers believed their child[ ]" and that "the proportion of mothers believing the report decreased to only 55.6% when the offender was a step-father or live-in partner" but not reporting what percentage of the accusations against biological fathers involved ex-spouses"). In the Tufts study, the family was not intact at the time of the abuse in 45% of the cases in which the natural fathers were accused. *See* TUFTS, *supra*. In an additional unspecified percentage of cases abuse was not revealed until after a divorce was complete. *See id.* Everson and his colleagues suggest that their findings reflect the fact that a mother is not as loving toward, or dependent on, an ex-husband as on a current mate. They point out, however, that none of the boyfriends in their sample admitted the abuse, whereas one-third of the biological fathers did so. *See* Everson et al., *supra*, at 205.

289 *See* KATHLEEN COULBORN FALLER, CHILD SEXUAL ABUSE: AN INTERDISCIPLINARY MANUAL FOR DIAGNOSIS, CASE MANAGEMENT, AND TREATMENT 43, 44 & tbl.2.27 (1988) (reporting that when rating 147 mothers of sexually abused children on scale from very protective to very unprotective—"[u]nprotective responses include disbelieving the child, blaming the child, and continuing to expose the child to risky situations after revelation of the sexual abuse"—53.1% rated as very unprotective, as somewhat unprotective, or as having switched from protective to unprotective); TUFTS, *supra* note 288, at 193, 194 tbl.7-1 (reporting that, as of the time of early treatment, 56% of the mothers of sexually abused children in the Family Crisis Program had failed to be consistently reassuring and supportive (18% not at all, 38% to some extent), 30% had reacted punitively (15% to some extent and 15% consistently), and 22% had demanded that the offender leave when possible); Christine Adams-Tucker, *Proximate Effects of Sexual Abuse in Childhood: A Report on 28 Children*, 139 AM. J. PSYCHIATRY 1252, 1255 (1982) (reporting that 65% of the mothers of 26

Mothers are most likely to be skeptical of their children's allegations when the child first reveals the abuse.[290] Such skepticism increases the likelihood that the child's revelation of abuse will be short-lived.[291] Mothers have many of the same motivations as their children do to deny that abuse has occurred and are well aware of the social and economic disruption that such allegations can create.[292]

One might expect that the legal system weeds out abuse cases in which the mother is unsupportive of the child's claim, which would

---

sexually abused children seen at a guidance clinic were unsupportive, which the authors defined as "knowing about the molestation but doing nothing, taking no action until the child became symptomatic, allowing the molester to be alone with the child again, delicately asking the molester to obtain counseling, believing the molester's denial, and ostracizing and blaming the child"); De Jong, *supra* note 288, at 16 & tbl.1, 17 & tbl.2 (reporting that of 103 children seen two to three weeks after medical evaluation for sexual abuse, 31% had mothers who were nonsupportive, which meant that the mothers "believed that the abuse complaint was a lie, a misunderstanding, or primarily the child's fault," and 12% of the children were not pressing charges); Everson et al., *supra* note 288, at 200 (reporting that in substantiated cases of sexual abuse, "44% of the 84 mothers were categorized as providing consistent support during the period following disclosure of sexual abuse, 32% were classified as ambivalent or providing inconsistent support, and the remaining 24% were unsupportive or rejecting of their children"); Margaret H. Myer, *A New Look at Mothers of Incest Victims*, 3 J. SOC. WORK & HUM. SEXUALITY 47, 49-53 (1985) (noting that in a study of 43 mothers of sexually abused children, 44% took no action (4/43) or rejected their daughters and protected their mates (15/43), and another 26% (11/43) were ambivalent, but sided with their daughters); Sirles & Franke, *supra* note 288, at 133 (reporting that in a study of sexually abused children, 21.8% of mothers did not believe their children had in fact been abused).

290 *See* Myer, *supra* note 289, at 55 ("When a mother is told that her daughter has been sexually abused by her mate, the first reactions are often shock and denial. . . . With professional intervention, 33 out of 43 mothers studied were able in time to accept that the abuse had occurred.").

291 *See, e.g.*, State v. Jackson, 730 P.2d 1361, 1362 (Wash. Ct. App. 1986) (revealing that when a five-year-old told her mother that the boyfriend had "'stuck his finger in [her] butt'" the mother had said "'Well, he'd better not have . . . did he really?'" to which the victim laughingly replied "'no'").

292 *See* FALLER, *supra* note 289, at 42-43 (reporting that of a sample of mothers rated on four-point scale from very independent to very dependant—dependence defined in part by "economic independence"—67.6% rated as somewhat (29.5%) or very dependent (38.1%)); TUFTS, *supra* note 288, at 188 (reviewing anecdotal data regarding mothers' denial and attributing this denial to "public humiliation," disruption of the family, "divorce and loss of financial support," and "fear [of] retaliation"); De Jong, *supra* note 288, at 18 (stating that "internal factors include denial, guilt, frustration, anger, fear of repercussions, feelings of inadequacy, ignorance, previous behavior or emotional problems of the child, or general distrust of or reluctance to involve the police, child protective services, or other agencies in personal matters" and that "[e]xterual factors would include pressures by family members or friends to protect the abuser, specific economic pressures that might arise from loss of support from the abuser, and lack of support and responsiveness from the police and social agencies involved in the investigation"); Myer, *supra* note 289, at 53 (finding that mothers who rejected daughters and protected their mates "were extremely dependent on their partners for emotional and economic support, and they all feared and were dominated by them"); *id.* at 57 ("These women face extreme emotional and economic stresses. . . . [T]hey face financial dependency, criticism from family, community agencies and officials, and isolation. The stronger the action they take in protecting the children, the more vulnerable they become to stress, loneliness, and deprivation.").

mean that a child testifying in a dependency or criminal action is more likely to have a supportive (and potentially suggestive) mother. On the contrary, children who were asked to testify in juvenile court "had mothers who were significantly less supportive than children who were not required to testify at the juvenile court hearing."[293] If a mother is unsupportive, dependency court involvement is more likely because her unwillingness to take action to protect the child against further abuse necessitates state intervention.[294] Moreover, no evidence suggests that mothers are more likely to be supportive in criminal cases.[295] In some cases, mothers even take the stand against their own children in defense of the accused molester.[296]

Regardless of children's motivation to protect their loved ones, they are less likely to be susceptible to suggestion about people they know well. In the Sam Stone Study, researchers created a "stereotype" by telling children various stories with the same theme: Sam was "a clumsy and bumbling person."[297] Because Sam Stone had visited the classroom only once for two minutes, however, the children lacked any personal knowledge of Sam with which the suggestion could compete.[298] If reseachers had given children extensive experience with a dextrous Sam, then the children may have formed a stereotype of Sam as a careful person and would have been better able to resist suggestions that he was the wrongdoer.

Ceci and his colleague's experimental and theoretical work provides evidence for this position.[299] Their work demonstrates that children's stereotypical knowledge of superheroes interferes with their long-term recollection of stories in which the superheroes act inconsistently with their stereotypes.[300] For example, researchers told children that the Six Million Dollar Man was unable to carry a can of

---

293 Everson et al., *supra* note 288, at 202.

294 *See, e.g., In re* Megan B., 1 Cal. Rptr. 2d 177, 179 (Ct. App. 1992) (noting that the dependency court faulted the mother for failing to believe child's statement that she was sexually abused).

295 *See* Everson et al., *supra* note 288, at 202 ("Mean levels of maternal support within the 'prosecution pending' and the 'no prosecution' subgroups did not differ significantly.").

296 *See, e.g.*, People v. Ford, 488 N.E.2d 573, 576 (Ill. App. Ct. 1985) (noting that the mother testified in defense of her boyfriend whom her eight-year-old daughter accused of rape); Williams v. State, 427 So. 2d 100, 102 (Miss. 1983) (noting that when the mother heard her 11-year-old daughter claim that the mother's boyfriend had raped the daughter, she "'told [the child] that she was a [G-d] damn liar!,'" that the child "'hadn't been raped, wasn't nothing wrong with her,'" and that "'[s]he just needs her damn brains beat out, that's all that was wrong with her'" (quoting testimony at the trial)), *legislatively amended by* MISS. R. EVID. 1103.

297 Leichtman & Ceci, *supra* note 34, at 570.

298 *See id.*

299 *See* Stephen J. Ceci et al., *Children's Long-Term Memory for Information that Is Incongruous with Their Prior Knowledge*, 72 BRIT. J. PSYCHOL. 443, 449 (1981).

300 *See id.* at 449.

paint because it was too heavy, and three weeks later the researchers asked the children whether the Six Million Dollar Man had been depicted as strong.[301] The authors found that "when new information is clearly incongruous with a child's preconceptions, although immediate recall may be accurate, shifts or distortions will subsequently occur" such that subsequent recall will move toward the child's preconceptions.[302] Although researchers *told* the children that the Six Million Dollar Man behaved weakly, they came to *believe* that his performance was consistent with his stereotype—the preconceived notion that he was strong.

Analogously, if a child has extensive experience with a person, so that the person becomes a loved one, then the child would form a stereotype of the person as one who cares for the child and would not cause needless harm to the child. Thus, the child more likely would resist stories that the loved one had abused the child. As the child's experiences with the loved one increase in number, the child's positive stereotype of that person becomes more detailed and better organized. Ceci argues that memory performance in a domain improves as knowledge in that domain improves in quantity and in structure.[303] For example, a child remembers the actions of her favorite playmate better than the actions of a relatively unfamiliar child.[304] As memory improves, suggestibility decreases.[305] Therefore,

---

301 *See id.* at 445.

302 *Id.* at 449.

303 Reviewing the literature on memory development, Chi and Ceci discussed a number of studies supporting the contention that "developmental differences are reduced when the amount of knowledge is somehow either controlled or equated." Michelene T.H. Chi & Stephen J. Ceci, *Content Knowledge: Its Role, Representation, and Restructuring in Memory Development*, 20 ADVANCES CHILD DEV. & BEHAV. 91, 115 (1987). They cite one study by Ceci and Howe finding that age differences in recall were attenuated if one controlled for the amount of knowledge children had about the to-be-remembered words. *See id.* at 114 (citing Stephen J. Ceci & Michael J.A. Howe, *Semantic Knowledge as a Determinant of Developmental Differences in Recall*, 26 J. EXPERIMENTAL CHILD PSYCHOL. 230 (1978)). Ceci and Bruck note that one even can reverse developmental differences in memory ability if the younger children are more knowledgeable than the older children. *See* Ceci & Bruck, *supra* note 9, at 415.

304 *See* Lynne Baker-Ward et al., *The Effects of Involvement on Children's Memory for Events*, 5 COGNITIVE DEV. 55, 65 (1990) (stating that "children in the more familiar group recalled more of the other-performed events than children in the less familiar group").

305 *See* Ceci & Bruck, *supra* note 9, at 417 (arguing that "the quality and quantity of memory representations influence subsequent recall and susceptibility to suggestibility"). Research literature disagrees about whether weaker memories are more susceptible to one type of suggestibility in which the original memory is actually distorted by the misinformation, a process called memory impairment. *See* CECI & BRUCK, *supra* note 36, at 254 (citing Mark L. Howe, *Misleading Children's Story Recall: Forgetting and Reminiscence of the Facts*, 27 DEVELOPMENTAL PSYCHOL. 746 (1991); Maria S. Zaragoza, *Preschool Children's Susceptibility to Memory Impairment*, *in* THE SUGGESTIBILITY OF CHILDREN'S RECOLLECTIONS 27 (John Doris ed., 1991)). Suggestibility, however, can occur without distortion of the original memory. Even if the details of one's original memory remain intact, misinformation may supplement memory and make the choice between the original memory and the misinformation

children should be less vulnerable to suggestions about those with whom they are especially familiar.

3. *Fears in the Real World*

Ceci and Bruck argue that studies of actual abuse cases further support their claim that threats do not suppress disclosure.[306] They discuss the results of two samples of abused children—a clinical sample reported by Sauzier and a sample of criminal cases reported by Gray.

> When the offender used aggressive methods to gain the child's silence, children were equally likely to tell about the abuse immediately following the event or to never disclose the abuse at all. Moreover, two thirds of children who were threatened not to tell nevertheless did disclose the details of their victimization. Thus, threatened children appeared to disclose as often as children who were not threatened.[307]

One cannot study the effects of threats on disclosure, however, by examining only those children whom researchers have identified as abused. Studies of cases in which children ultimately revealed abuse are problematic because these studies exclude the very children for whom threats were most effective in suppressing reports of abuse. A

---

difficult. Zaragoza examined one type of suggestibility in which memory for the original event is "impaired." Zaragoza, *supra*, at 27. Her procedure does not test for suggestibility attributable to gap filling when one has forgotten, to supplementation of the original memory, or to difficulties in distinguishing between one's memory of the original event and one's memory of the suggested information (i.e., source monitoring difficulties). *See id.* at 28, 37. Similarly, Howe found that "trace strength [the strength of a memory] is directly related to the rate of forgetting . . . and the number of . . . misinformation-relevant intrusions, [but] it does not impair recall of the original story details." Howe, *supra*, at 760. Furthermore, some have criticized Howe's finding on the ground that "the fact that the overall effect of the misled/control condition manipulation was small to nonexistent suggests one reason why the potential variance explained by the interaction of the number of training trials [which affects memory strength] with the misled/control condition may have been so limited." Kathy Pezdek & Chantal Roe, *The Effect of Memory Trace Strength on Suggestibility*, 60 J. EXPERIMENTAL CHILD PSYCHOL. 116, 125 (1995).

Several studies examining suggestibility that do not separately test for memory impairment effects have found that suggestibility effects are larger when memory is weaker. *See, e.g.*, Goodman et al., *supra* note 23, at 262 (discussing Rudy and Goodman's findings that "age differences in suggestibility result at least in part from younger children having weaker memories than older children for certain types of information"); Pezdek & Roe, *supra*, at 124, 125 (concluding that "stronger memories are more likely to resist suggestibility than weaker memories"); Amye Warren et al., *Inducing Resistance to Suggestibility in Children*, 15 LAW & HUM. BEHAV. 273, 282 (1991) ("If one has a weak memory trace, then either negative feedback or misleading information alone may produce the uncertainty that leads to either acquiescence or inconsistencies in recall. One with a stronger memory trace, on the other hand, may not succumb to doubt until the two are combined.").

306 *See* CECI & BRUCK, *supra* note 36, at 301.

307 *Id.*; *see also id.* at 35 ("Some experts state that children do not disclose because of explicit threats made by the perpetrators. The available evidence does not support this assertion.").

child's statement is the most common means of detecting abuse.[308] If threats in fact suppress reporting, then the percentage of allegedly abused children who report having been threatened will underestimate the actual percentage of abused children who are threatened. Moreover, the relation between threats and willingness to report among children known to have suffered abuse may not reflect the actual relation between threats and reporting among all abused children.

If threats reduce the willingness of children to report abuse, but do not eliminate reporting altogether, then one could examine the relation between threats and the time at which children ultimately reveal abuse. However, the process by which investigators substantiate reports complicates this analysis. Cases involving reluctant children are less likely to become substantiated. Therefore, even partially effective threats will have the tendency to exclude temporarily silenced children from studies of substantiated cases of abuse.

The fact that substantiated cases of abuse do not represent all cases of abuse explains the apparent paradox that abused children are reluctant to disclose their abuse while most substantiated cases of abuse involve children who have disclosed.[309] Problems of representativeness become more serious as one moves from social services substantiation to juvenile court involvement to criminal court involvement. The more reluctant or resistant the child, the less likely that the case will survive higher burdens of proof. Ceci and Bruck recognized this point, noting that children in "clinical" samples of abuse are probably less forthcoming about their abuse than children in "forensic" samples.[310] Even less forthcoming than either of those

---

308 *See* DEBRA WHITCOMB ET AL., THE CHILD VICTIM AS A WITNESS 88, 92 (1994) (reviewing a sample of 431 sexual abuse cases referred for prosecution and concluding that "[i]n the vast majority of cases, the child victim disclosed the abuse"); Haskett et al., *supra* note 241 (noting that the most important factor in the substantiation process is whether the child discloses abuse).

309 *See, e.g.*, Leslie Biron Campis et al., *Developmental Differences in Detection and Disclosure of Sexual Abuse*, 32 J. AM. ACAD. CHILD & ADOLESCENT PSYCHIATRY 920, 923 (1993) (noting that most cases in their sample were purposeful disclosures whereas population surveys show purposeful disclosure is rare). *But see* Ceci et al., *supra* note 33, at 503 ("Lyon cannot have it both ways: either 86% of children suspected of having been abused disclose quickly and are interviewed repeatedly about the details of their quick disclosures, or they deny having been abused during the initial interview and require multiple interviews.").

310 CECI & BRUCK, *supra* note 36, at 35. Ceci and Bruck note:

> Children in forensic samples may be those who readily disclose, whereas children in clinical samples who delay making disclosures may not go through the criminal system as readily; these may be the children for whom it is difficult to extract a report, and thus they are brought by adults for treatment. Finally, these studies provide no information on the number of children or the profiles of children who never disclose.

Id.

are the abused children who are not included in either sample because they fail altogether to disclose their abuse.

Two lines of research substantiate the underreporting problems. First, surveys of adults consistently find that a large percentage of adults now willing to talk about their abuse never revealed it as children,[311] and even fewer of these cases were ever reported to the police or resulted in prosecutions.[312] Second, studies of children who show medical evidence of sexual abuse find that from thirty-five percent to fifty percent of these children fail to disclose their abuse.[313]

Even if one overlooks the difficulties of interpreting data on children who ultimately reveal abuse, such data fail to support Ceci and Bruck's claims. Ceci and Bruck cite Sauzier's study[314] for the proposition that when the abuser used "aggressive methods to gain the child's compliance to keep the secret, children were equally likely to tell about the abuse immediately following the event or to never disclose the abuse."[315] In the cited study, however, Sauzier referred to cases in which the abuser used aggression to abuse the child, not to elicit secrecy.[316] Moreover, even if one assumes that aggressive abusers always aggressively threaten children not to reveal, the fact that an equal number of these children disclose as fail to disclose does not resolve the question of whether aggression reduces disclosure. One must compare this disclosure rate to the disclosure rate for children who have not been aggressively abused or threatened.

Ceci and Bruck do not mention the explicit comparison that Sauzier performed between cases involving aggressive abuse and those involving abuse accomplished through manipulation or threats:[317]

---

311 *See* Lyon, *supra* note 240 (manuscript at 6) (reviewing research and indicating that "[r]ates of non-disclosure among women run from 33% to 92%; among men from 42% to 85%").

312 *See id.* (manuscript at 11) (reviewing research).

313 *See supra* text accompanying notes 243-46. Researchers studying cases in which the evidence for sexual abuse was less definitive have reported similar results. *See* Rosemary S. Hunter et al., *Sexually Abused Children: Identifying Masked Presentations in a Medical Setting*, 9 CHILD ABUSE & NEGLECT 17, 21 (1985) (describing a study in which 50 children who were seen in the hospital and whose initial allegations did not involve abuse, but who were ultimately reported to social services as sexually abused and concluding "in 26% [of the cases] the abuser could not be established at the time of report to social services").

314 *See* Maria Sauzier, *Disclosure of Child Sexual Abuse: For Better or for Worse*, 12 PSYCHIATRIC CLINICS N. AM. 455 (1989).

315 CECI & BRUCK, *supra* note 36, at 35.

316 *See* Sauzier, *supra* note 314, at 459 (referring to "strategies for gaining the child's compliance").

317 Although Sauzier's report states that aggression was the most common strategy, *see id.*, the study upon which Sauzier bases her report, the Tufts study, states that "most of the offenders (83 percent) used manipulation," TUFTS, *supra* note 288, at 87. It looks as if the table describing the effect of the offenders' strategies reverses the numbers using each strategy. *Compare id.* at 88 (listing the number of cases of manipulation as 130 and of

> The offenders' strategies for gaining the child's compliance were also related to disclosure: Aggressive methods were more likely to evoke either immediate reporting (39 per cent) or failure to ever tell (43 per cent). . . . Most children subjected to intercourse with aggression never revealed. When the strategy used relied on manipulation, only 25 per cent of children reported the abuse immediately. Threats also seemed to prevent children from telling immediately (only 23 per cent did).[318]

Sauzier also investigated the relationship between fearfulness and reluctance to disclose among abused children and concluded that

> children who failed to reveal more serious abuse had the highest fear scores. They described the fear of losing the affection and goodwill of the offender; fear of the consequences of telling (being blamed or punished for the abuse by the non-offending parent); fear of being harmed; and fear of retaliation against someone in their family.[319]

Firm conclusions based on Sauzier's data are problematic—the sample may not be representative of abused children generally, and the differences may not be statistically significant.[320] Nevertheless, one fairly can question the assertion "that the likelihood of disclosure was unrelated to claims of threats by the offender."[321]

Ceci and Bruck also cite Gray's study of criminal sexual abuse prosecutions.[322] Gray's study found that children who were threatened by their abuser were just as likely as children who were not threatened to disclose abuse before questioning.[323] However, children against whom threats are most effective are least likely to appear among cases prosecuted in criminal court. Moreover, unlike Sauzier, Gray did not examine whether threatened children delayed longer than nonthreatened children before revealing abuse. A Canadian study of 135 children whose cases the Government prosecuted in criminal court found that "threats were far more common" among those children who delayed reporting their abuse.[324]

---

aggression as 54), *with id.* at 100 (listing the number of cases exactly backwards with 54 cases of manipulation and 130 cases of aggression).

318 Sauzier, *supra* note 314, at 459.

319 *Id.* at 460; *see also supra* note 266 (providing the most common explanations for why abuse victims failed to reveal abuse).

320 Sauzier did not test any differences for statistical significance. *See* Sauzier, *supra* note 314.

321 CECI & BRUCK, *supra* note 36, at 301.

322 *See* GRAY, *supra* note 222, at 31 (studying 670 children alleged to be victims of sexual abuse).

323 *See* CECI & BRUCK, *supra* note 36, at 35.

324 LOUISE DEZWIREK SAS & ALISON HATCH CUNNINGHAM, TIPPING THE BALANCE TO TELL THE SECRET: PUBLIC DISCOVERY OF CHILD SEXUAL ABUSE 122 (1995). The study noted that "overt threats were not necessary if the man had a history of violence within the

The existence of threats partially explains why interviewers may feel compelled to ask leading questions when interviewing children about abuse. Research on keeping secrets suggests that children are more likely to reveal abuse as questioning becomes more direct.[325] However, leading questions do not always elicit disclosures of wrongdoing. Moreover, these questions may elicit false disclosures. Inevitably, one faces a tradeoff between false accusations and false denials.

B. Embarrassment

For many children, sexual abuse is embarrassing. In their study of young girls' memories of genital touching by a pediatrician,[326] Saywitz, Goodman, and their colleagues attributed much of the underreporting in free recall to embarrassment rather than to memory retrieval difficulties. If memory were the culprit, then one would expect the seven-year-olds to recall genital touch more accurately than the five-year-olds. Instead, the study detected the opposite, suggesting that the seven-year-old girls were more aware of the embarrassing nature of genital touch, even when performed by a doctor as part of a parentally sanctioned examination. More generally, touching by unfamiliar adults appears to evoke some discomfort among children. Several studies have found that false negative responses occur more often when children are asked about touching rather than about other actions by strangers.[327]

---

home," which would attenuate any apparent relation between threats and delayed disclosure. *Id.*

325 *See, e.g.*, Wilson & Pipe, *supra* note 261, at 68 (noting that children were more forthcoming about the inkspill when directly questioned about it than when asked for free recall or cued recall).

326 *See* Saywitz et al., *supra* note 59.

327 *See* Goodman et al., *supra* note 23, at 262 (discussing Rudy & Goodman, *supra* note 52, regarding "abuse" questions and indicating that "[t]he more common error was to omit actions that did occur" and that "[e]ven then, virtually all of the omission errors made by the 7-year-olds were in response to a specific subset of questions—those concerning touching"); Michael R. Leippe et al., *Eyewitness Memory for a Touching Experience: Accuracy Differences Between Child and Adult Witnesses*, 76 J. APPLIED PSYCHOL. 367, 375 (1991) ("[W]ith regard to the most salient action in the skin-test situation—the toucher's touching of the subject—the memory errors of 5- to 6-year-olds were primarily restricted to the failure to report touches that did occur rather than the reporting of touches that did not occur."); Douglas P. Peters, *The Influence of Stress and Arousal on the Child Witness*, *in* THE SUGGESTIBILITY OF CHILDREN'S RECOLLECTIONS, *supra* note 305, at 60, 63-65 (explaining that children in the touched group "were vigorously rubbed on the head until they attempted to avoid the rubbing by flinching their heads away or verbally protesting" and that "for the 34 children in the [t]ouched group a significant number (56%) made more false negative responses (did not report that the stranger had in fact touched [rubbed] their heads) than correct recall (32%) or false positives (12%) in which other parts of the body were identified"). Discussing Peters's results, Ceci and his colleagues suggested that "the children may have been motivated not to reveal having been touched due to a fear of embarrassment." Ceci et al., *supra* note 48, at 124.

If children are embarrassed to talk about innocuous touching, they are likely less inclined to disclose touching they believe will evoke disapproval. Goodman and others consistently have found that false affirmations are less common when researchers ask abuse-related questions.[328] Moreover, Goodman observed that children's demeanor often changed when interviewers asked abuse-related questions, such as whether an unfamiliar male kissed the child or took the child's clothes off. Children giggled, smiled, looked surprised, or exhibited disgust.[329]

Ceci and Bruck point out that "the literature clearly does not support the strong view that bodily acts are impervious to distortion"[330] and cite several studies in which substantial numbers of children falsely reported bodily touch.[331] Yet Ceci and Bruck do not deny that embarrassment both suppresses true reports of abuse and decreases the likelihood of false reports of actions that might suggest abuse.[332] Indeed, Ceci and Bruck have cited research demonstrating that adults report having failed to disclose abuse as children due to embarrassment and self-blame.[333]

---

328 *See, e.g.*, Rudy & Goodman, *supra* note 52, at 535.

329 *See* Goodman et al., *supra* note 23, at 266. Goodman and colleagues, discussing the Rudy and Goodman study, stated:

> We also noticed that children's demeanor changed once we began to ask the abuse questions. Many showed signs of embarrassment by giggling or smiling. Others looked surprised. Some covered their eyes with their hands, puckered up their faces in disgust, asked in disbelief if we would repeat the question, or, if their parent was in the room during the questioning, glanced over at him or her in an act of "social referencing," with a look of "good grief!" on their faces. We scored the children's nonverbal responses to three of the most blatant abuse questions: "Did he kiss you?" "He took your clothes off, didn't he?" and "Did he hit you?" in comparison to the questions preceding that line of questioning, which mostly concerned the confederates' appearance. The children showed significant increases in smiling and surprise as soon as the abuse questions began.

Id.

330 CECI & BRUCK, *supra* note 36, at 234.

331 Several of these studies involve high rates of false responding by three-year-olds, *see id.*, which raises the issue of the age at which children exhibit embarrassment about touching. Younger children likely do not feel such embarrassment.

332 *See* Ceci et al., *supra* note 33, at 506 ("[T]ruly abused children are often unlikely to disclose sexual abuse out of a sense of embarassement or fear . . . .").

333 *See* CECI & BRUCK, *supra* note 36, at 203 (citing Donna Della Femina et al., *Child Abuse: Adolescent Records vs. Adult Recall*, 14 CHILD ABUSE & NEGLECT 227, 228-29 (1990)). Of a group of 69 subjects interviewed both as adolescents and, nine years later, as adults, "26 gave responses regarding [physical] abuse that were discrepant with information gathered when they were adolescents." Femina et al., *supra*, at 228. Eleven of these adults agreed to a "Clarification Interview" during which they provided explanations for these inconsistencies:

> [A]ll 11 subjects with discrepant data who were reinterviewed had, as far as could be ascertained, been abused. Reasons for denial, whether in adolescence or adulthood, included embarrassment, a wish to protect parents, a sense of having deserved the abuse, a conscious wish to forget the past, and a lack of rapport with the interviewer.

Ceci and Bruck openly acknowledge the role of embarrassment in reducing false claims of abuse in Goodman's and Saywitz's research.[334] Ceci and Bruck treat this fact, however, as if it were a criticism of the research, rather than a point in favor of its ecological validity.[335] The authors accuse Goodman and Saywitz of "tilting the motivational structure toward truthful reporting" because "if children in these earlier studies were to distort what they had witnessed, and claim to have been sexually touched when they were not, this could be expected to result in embarrassment."[336] If one is interested in whether children falsely will accuse adults of sexual abuse, it is hard to understand why asking questions about sexual abuse, as opposed to less embarrassing actions, is "tilting the motivational structure."[337] From the perspective of an abuse investigator looking to the literature for advice, the tilt seems to occur in studies like Sam Stone.

Saywitz and colleagues failed to provide children who had not been touched with strong motives falsely to disclose abuse,[338] and Ceci and Bruck argue that a desire to avoid embarrassment motivates false allegations.[339] Ceci and Bruck overlook two issues: first, whether interviewers in fact embarrass children into acknowledging abuse; and second, whether embarrassment is equally effective as an inducement to reveal and as a motivator to conceal.

How does one embarrass a child into alleging abuse? Ceci and Bruck suggest a question like "He kissed you because he loves you, didn't he?"[340] One can only describe this question as bizarre.[341] Young children, unaware of societal disapproval of sexual abuse, may misinterpret abusive acts as appropriate affection, at least when coercion, secrecy, or discomfort does not accompany the abuse. Most consider such a reaction detrimental to the child, in part because the child may behave sexually in a manner that evokes disapproval and puts the child at risk of future abuse.[342] An interviewer who "embar-

---

Id. at 229; *see also* Johnson & Shrier, *supra* note 266, at 374 (noting that some adolescent males had not reported childhood abuse because "they wanted to forget about the incident, wanted to protect the assailant, or were afraid of the reactions of their peers and family members").

334 *See* CECI & BRUCK, *supra* note 36, at 265.

335 *See id.*

336 *Id.*

337 *Id.*

338 *See* Saywitz et al., *supra* note 59, at 691 (stating that "[c]hildren in the nongenital condition had no motive to distort their reports").

339 *See* CECI & BRUCK, *supra* note 36, at 265-66.

340 *Id.* (internal quotation marks omitted).

341 Ceci and Bruck may have seen this question somewhere, given the large number of interviews they reviewed. Unfortunately, they do not provide any evidence of its use.

342 *See, e.g.*, David Finkelhor, *The Trauma of Child Sexual Abuse: Two Models*, 2 J. INTERPERSONAL VIOLENCE 348, 355-56 (1987) (arguing that sexual abuse shapes children's sexuality in "developmentally inappropriate and interpersonally dysfunctional ways," in part

rasses" a child into reporting abuse in the manner Ceci and Bruck imagine is not only suggestive, but blatantly inappropriate.

To demonstrate the efficacy of embarrassment in creating false allegations, the authors cite a study that Ceci and his colleagues conducted of four three-year-old children whose parents gave them baths.[343] Parents kissed two of the children while the children were in the bathtub and did not kiss the other two children. Interviewers told the two children the parents had kissed that it was naughty to be kissed while naked and then asked the children leading questions designed to embarrass them into a false denial. The two children whom the parents kissed both denied having been kissed.[344] Interviewers told the two children the parents had not kissed that parents who loved their children often kissed them while giving them a bath and then asked a leading question designed to embarrass them into making a false allegation. One of the children whom the parents had not kissed falsely admitted having been kissed, although she later reversed her story when a parent interviewed her alone.[345]

One may be tempted to interpret the fact that the false negative rate is twice as high as the false positive rate to mean that embarrassment suppresses true reports more than it elicits false accusations. These results prove little, however, because the sample size was only four. The fact that the sample size was not larger is itself significant. McGough notes that the researchers aborted the study and describes the reaction of the two children whom researchers told that kissing was naughty: "Before the planned reassurance could be completed—that it was not bad if the touching were by a parent or close relative—both children displayed high anxiety and began to cry silent tears."[346] It is instructive to explore why the children whom researchers told that loving parents kiss their children did not cry. Possibly, three-year-old children happily would acknowledge that their parents kissed them in the tub without any embarrassment induction, and parents may in fact routinely do so (though they did not on this particular occasion). If anything, the study provides a compelling anecdote that children may be reluctant to affirm events that they believe are naughty.

Ceci and Bruck argue that embarrassment is not the sole motivator of children in sexual abuse intervews and cite numerous other

---

because "[c]hildren become confused and acquire outright misconceptions about sexual behavior and sexual morality as a result of things that offenders tell them").

343 *See* CECI & BRUCK, *supra* note 36, at 264-65.

344 *See id.*

345 *See id.* at 265.

346 MCGOUGH, *supra* note 13, at 289 n.11 (citing Stephen J. Ceci & Michael P. Toglia, Presentation at the Biennial Meeting of the American Psychology and Law Society (Mar. 16, 1990)).

"dominant motivations" that interviewers may bring to bear on children.[347] This argument, however, sidesteps the simple assertion that embarrassment reduces the likelihood that children falsely will affirm sexual touching. Research that fails to ask about sexual touching exaggerates children's susceptibility to suggestion in sexual abuse cases. If a researcher incorporates the other "dominant motivations," but does not account for embarrassment, then he arguably has "tilt[ed] the motivational structure"[348] in favor of his desired outcome.

The argument that other motivations may override embarrassment also ignores the fact that some motivations are more powerful than others. In Ceci and Bruck's view, embarrassment and fear of reprisal appear to be the most powerful motives inducing children to make false statements.[349] When one asks which factor is more likely to induce fear and embarrassment in children, their awareness of societal attitudes toward sexuality or interviewer influence, the fact that an interviewer must be heavy-handed indeed to embarrass and frighten a child into falsely acknowledging abuse proves particularly striking.[350]

C. Recall Versus Recognition

Most laypeople have at least some familiarity with the distinction between recall and recognition, a distinction that memory researchers have found important for many years. In the classic memory experiment, researchers ask a subject to memorize a list of words. Subsequently, the researcher tests the subject's memory of the words. If the test were for recall, the experimenter would ask something like, "What were the words on the list you memorized?" If the test were for recognition, the experimenter would recite words and ask the subject whether she remembered seeing them on the list.

In a recall test, the subject must generate the words herself, whereas in a recognition test, she merely identifies the words the experimenter generates. Therefore, as one moves from recall to recognition, the amount of detail in the question increases, and the required response moves from a narrative response to a yes-or-no response. One also may give the subject a test that lies between recall and recognition. In this test, the experimenter gives the subject some

---

347 Ceci et al., *supra* note 33, at 504.

348 CECI & BRUCK, *supra* note 36, at 265.

349 *See* Ceci & Bruck, *supra* note 9, at 427.

350 Children can overcome embarrassment. Ceci and his colleagues cite transcripts of interviews in which young children "sometimes seemed gleeful as they recounted their alleged sexual molestation." Ceci et al., *supra* note 33, at 506. Their acknowledgment that such glee "may not characterize the vast majority of child sexual abuse cases," *id.*, matches my own experience in interviewing young children. Although I am not aware of any corroborating research, in my experience children's humor regarding bodily functions tends to involve urination and defecation rather than sexual acts.

help in generating the desired information. The experimenter might provide hints—or cues—regarding the type of words that were on the list (e.g., "What animals were on the list?"). One would call such questions "cued-recall."

Recall is more difficult than recognition.[351] Recall proves particularly difficult for young children. Ceci and Bruck note that "age differences in recognition memory are far less pronounced than age differences in free recall, and at times these are nonexistent."[352] Moreover, cues can reduce age differences in recall performance.[353]

Recognition tests, however, do impose certain costs. Although experimenters can elicit more information through recognition questions, the number of errors also increase with the use of such questions. Recognition questions tend to be leading because such questions contain information the child previously has not disclosed and because the child may assume that the interviewer desires a "yes" response. In the context of interviewing children about abuse, a false "yes" presents two dangers: first, the child may simply respond "yes," leading the interviewer to believe something untrue; and second, the child actually may come to believe that the answer is "yes," either because the interviewer implicitly suggests this answer or because the child subsequently recalls her answer better than the original event.

In interviewing terminology, recall is analogous to open-ended questions, and recognition is analogous to direct or leading questions. Consider the continuum of questions Bruck and Ceci outlined. Open-ended questions include "Can you tell me about what happens at naptime?"[354] Specific questions include "Who is in the room at naptime?" and "Do people do anything special at naptime?"[355] Leading questions (other researchers would call these "direct" or "spe-

---

351 Indeed, developmentalists have demonstrated that children understand this fact about memory by the time they are five years old. *See* James Ramsey Speer & John H. Flavell, *Young Children's Knowledge of the Relative Difficulty of Recognition and Recall Memory Tasks*, 15 DEVELOPMENTAL PSYCHOL. 214, 217 (1979).

352 Ceci & Bruck, *supra* note 9, at 404 (emphasis omitted).

353 *See* Stephen J. Ceci & Michael J.A. Howe, *Age-Related Differences in Free Recall as a Function of Retrieval Flexibility*, 26 J. EXPERIMENTAL CHILD PSYCHOL. 432, 435-36 & tbl.2 (1978). This study examined a sample of 72 children, 24 each in nursery school, second grade, and fifth grade. *See id.* at 434. Children were presented with 25 drawings in an incidental learning task in which researchers trained children to group pictures thematically and taxonomically. *See id.* Subsequently, interviewers conducted both cued recall and free recall tests with each child. *See id.* at 434-35. "Although 10-year-olds' free recall was nearly 50% greater than 4-year-olds [sic], their cued recall was only 8% greater." *Id.* at 435-36.

354 Bruck & Ceci, *supra* note 39, at 306 (internal quotation marks omitted).

355 *Id.* (internal quotation marks omitted).

cific")[356] include "Did anything scary happen at naptime?" and "Did anyone ever touch you in a bad place at naptime?"[357]

Researchers disagree about whether one necessarily must ask children recognition questions in abuse investigations. On the one hand, several studies have found that when interviewers supplement open-ended questions with recognition questions, children produce more information about experienced events. In particular, preschool children exhibit the greatest increase in memory performance when interviewers use this technique. For example, Ornstein and his colleagues examined three- and six-year-old children's memories of pediatric examinations.[358] The interviews moved from open-ended to recognition ("yes/no") questions. Three-year-old children produced three to four times as much information when one added their responses to yes/no questions to their answers to open-ended questions.[359] Adding yes/no questions to open-ended responses more than doubled the amount of information six-year-old children produced.[360] As Ceci and his colleagnes have recognized, "it is terribly difficult to elicit recall from 3-year-olds (our youngest age group), and often the sort of prompting that is necessary to get at the complete contents of their memory is similar to actually using a recognition procedure."[361]

On the other hand, Lamb and his colleagues repeatedly have demonstrated that children provide more details and longer responses in abuse investigations to invitation or recall questions than to

---

356 The difficulties in defining what questions should be characterized as "leading" is discussed *supra* in notes 178-85 and accompanying text.

357 Bruck & Ceci, *supra* note 39, at 306 (internal quotation marks omitted).

358 *See* Peter A. Ornstein et al., *Children's Memory for a Personally Experienced Event: Implications for Testimony*, 6 APPLIED COGNITIVE PSYCHOL. 49 (1992).

359 *See id.* at 55 tbl.3. In a similar study, Lynne Baker-Ward and colleagues added information elicited through yes/no questions during an interview immediately after the examination and found that such information nearly tripled the memory performance of the three-year-olds, nearly doubled the performance of the five-year-olds, and increased the proportion of features remembered by the seven-year-olds by 50%. *See* Lynne Baker-Ward et al., *Young Children's Long-Term Retention of a Pediatric Examination*, 64 CHILD DEV. 1519, 1524-26 (1993). Difference between recall and yes/no performance was even larger after one- to six-week delays. *See id.*

360 *See* Ornstein et al., *supra* note 358, at 55 tbl.3.

361 Stephen J. Ceci et al., *On Remembering . . . More or Less: A Trace Strength Interpretation of Developmental Differences in Suggestibility*, 117 J. EXPERIMENTAL PSYCHOL.: GEN. 201, 202 (1988); *see also* CECI & BRUCK, *supra* note 36, at 71 (summarizing Ornstein and colleagues' research and noting that "[t]he 3-year-olds were particularly noteworthy for the lack of information that they provided to the open-ended questions, thus forcing the interviewer to ask a large number of specific yes/no questions in order to obtain a full report about the visit"); Ornstein et al., *supra* note 358, at 58 (noting that "it was necessary to rely more fully on yes-no, specific probes when dealing with the 3-year-olds, because these children generated relatively little information in response to the open-ended questions").

focused or recognition questions.[362] Poole and Lamb concluded that "[a]lthough *individual* specific questions elicit much less information than individual open-ended prompts, researchers consistently have shown that children provide fewer details in response to open-ended questions than in response to a series of specific questions."[363] Hence, although open-ended questions may produce a greater yield per question, specific questions produce a greater overall yield. Unfortunately, one cannot use Lamb's data to determine whether it is possible to elicit most or all of the details of abuse by asking more open-ended questions because the interviewers in the Lamb studies ask so few open-ended questions.[364] Furthermore, the work of Lamb and his colleagnes suggests that open-ended questions produce their greatest advantage with older children, but whether preschool children also would benefit is unknown.[365]

Because of the risk of misleading children by asking direct questions, both researchers and practitioners recommend that interviewers begin with open-ended questions and move to direct, specific, or leading questions only if the child is nonresponsive.[366] A questioner's

---

362 *See* Hershkowitz et al., *supra* note 174, at 172-73; Lamb et al., *supra* note 173, at 1255-56; Lamb et al., *supra* note 167, at 633-34; Kathleen J. Sternberg et al., *Effects of Introductory Style on Children's Abilities to Describe Experiences of Sexual Abuse*, 21 CHILD ABUSE & NEGLECT 1133, 1139 (1997); Sternberg et al., *supra* note 172, at 447-48.

363 DEBRA A. POOLE & MICHAEL E. LAMB, INVESTIGATIVE INTERVIEWS OF CHILDREN 52 (1998) (emphasis added) (citations omitted).

364 In one study, Sternberg and her colleagues manipulated whether investigators used an open-ended introductory style or a direct introductory style. *See* Sternberg et al., *supra* note 362. They found that children responded to the first abuse-related question at greater length and in more detail when the investigator used the open-ended introductory style. *See id.* at 1140. Investigators using the open-ended introductory style, however, did not continue to ask open-ended questions, and therefore one could not determine whether an interview in which open-ended questions predominated was superior. *See id.* ("Interviewers who began the interviews using the open-ended introductory protocol did not use more open-ended utterances in their interviews than those using the direct introductory protocol.").

365 *See* Lamb et al., *supra* note 173, at 1257 ("The available sample was too small for us to examine age differences, although it is possible that different results would have been obtained had the children been younger."); Lamb et al., *supra* note 167, at 635 ("Because we limited this study to children between 5 and 11 years of age, we do not know whether similar results would have been obtained had the study been focused on preschool-aged children."); Sternberg et al., *supra* note 362, at 1137 (explaining that 11 of 51 or 22% of children interviewed were four to six years of age, and the youngest was four and one-half); *id.* at 1139-40 (explaining that children older than eight exhibited a larger increase in details than children under eight when interviewers used an open-ended introduction rather than a direct introduction and noting that details increased by 180% for older children and 42% for younger children); Sternberg et al., *supra* note 172, at 442 (involving interviewees four to twelve years of age); *id.* at 447 ("[T]he superiority of open-ended over focused questions was greater for older [8-11 and older] than for younger children [8-11 and younger].").

366 *See, e.g.*, AMERICAN PROSECUTORS RESEARCH INST., NAT'L CTR. FOR PROSECUTION OF CHILD ABUSE, INVESTIGATION AND PROSECUTION OF CHILD ABUSE at II-8 to II-9 (1987) (recommending that interviewers avoid leading questions); HOME OFFICE & DEP'T OF HEALTH,

judgment regarding the extent to which direct, specific, or leading questioning is appropriate depends upon the questioner's estimation of the tradeoff between the additional amount of true information elicited and the increased risk of false affirmations.

Results from the Inoculation and the Sam Stone Studies confirm the risks of increasingly focused questions. During the final interview in the Inoculation Study, the experimenter first asked the child to recall everything she remembered the time she had her shot.[367] The experimenter then showed the child pictures of the research assistant and pediatrician, in turn, and asked the child to describe what each person had done during the examination.[368] Finally, the experimenter asked the child specific questions regarding who performed each procedure.[369] The largest percentage of false allegations came in response to the specific questions, the next largest came in response to the probe question regarding each persons' actions, and the smallest percentage came in response to the request for free recall.[370] In the Sam Stone Study, although a substantial number of the children subjected to suggestive interviews referred to nonexistent events in their response to a free narrative question, the percentage of false reports increased in response to a probe question.[371]

These two studies failed to consider the extent to which a free recall request might limit the amount of information interviewers can elicit from young children. Both studies focused on children's errors, and therefore they reported only the percentage of incorrect responses to the various types of questions. Ceci emphasized that "when there was no attempt by the interviewers in the Sam Stone Study to

---

MEMORANDUM OF GOOD PRACTICE: ON VIDEO RECORDED INTERVIEWS WITH CHILD WITNESSES FOR CRIMINAL PROCEEDINGS 17 (1992) (recommending that interviewers begin with open-ended questions); R. Edward Geiselman et al., *Effects of Cognitive Questioning Techniques on Children's Recall Performance, in* CHILD VICTIMS, CHILD WITNESSES 71, 78-79 (Gail S. Goodman & Bette L. Bottoms eds., 1993) (detailing elements of "cognitive interview," which includes asking for narrative report before asking specific questions); John C. Yuille et al., *Interviewing Chidren in Sexual Abuse Cases, in* CHILD VICTIMS, CHILD WITNESSES, *supra*, at 95, 99 (recommending that interviewers "begin with the most open, least leading form of questioning and proceed to more specific forms of questioning as circumstances require").

367 *See* Bruck et al., *I Hardly Cried, supra* note 34, at 205.

368 *See id.*

369 *See id.*

370 *See id.* Eleven percent (9/85) of the total false allegations were in response to an open-ended inquiry about the examination, 38% (32/85) were in response to "What did [the RA] do?," and 48% (41/85) were in response to "Who gave you your shot?," indicating that errors increase with the number of specific questions asked. *Id.* (internal quotation marks omitted).

371 *See* Leichtman & Ceci, *supra* note 34, at 572-73. In the Sam Stone Study, for the free narrative question the researcher asked the children to "tell [her] everything that happened?" *Id.* (internal quotation marks omitted). For the probe question, the researcher asked the children if they had "heard something" about a book and if they knew anything about that. *Id.* (internal quotation marks omitted).

mislead them, even 3-year-olds recalled large amounts of information accurately."[372] The Sam Stone Study, however, contained little accurate information for the children to report. As Ceci and Bruck have noted, "the 2-minute visit of Sam to the classroom is not a significant event. . . . There really was *no* event."[373]

In their brief before the New Jersey Supreme Court, Bruck and Ceci refrained from openly endorsing a position regarding direct or leading questions. They reviewed guidelines that allow for specific questions and indicated that "[*s*]*ome* interviewers advocate the use of leading questions as a *last resort,* if the child provides no information in the interview."[374] An example of a "last resort" question that only "some" interviewers advocate ever using is "Did anyone ever touch you in a bad place at naptime?"[375] This type of question, unlike the highly coercive and suggestive questions documented in the *Michaels* case, fails to name a suspected offender and fails to specify what being touched in a "bad place" means.[376]

The Ceci and Bruck brief also failed to discuss the research findings that young children have difficulty generating responses to free recall questions. Nor did the New Jersey court acknowledge this difficulty in establishing the standard for taint hearings. The court held that the first factor the trial courts must consider in determining whether to hold a hearing is the "absence of spontaneous recall" in response to an interviewers' questions.[377] If the defendant provides "some evidence" that pretrial questioning influenced a child's report, the court will bar that child from testifying unless the state demon-

---

372 Ceci, *supra* note 32, at 48; *see also* Leichtman & Ceci, *supra* note 34, at 571-72 ("[The youngest] children's reports usually included accurate accounts of actual information; they often were able to recall Sam Stone's limited activities on the day he visited, for example, that he walked around the housekeeping section of the classroom, that he greeted the children pleasantly, or that he waved goodbye.").

373 CECI & BRUCK, *supra* note 36, at 133; *see also* Leichtman & Ceci, *supra* note 34, at 575 ("[Q]uite accurately, . . . [one child] asserted that Sam Stone had come into the classroom and said hello and looked around, but that 'nothing happened.'").

374 Bruck & Ceci, *supra* note 39, at 306 (emphasis added).

375 *Id.* (internal quotation marks omitted) (indicating that some interviewers use such questions if the child provides no information in the interview).

376 It is not clear, for instance, whether "bad place" refers to a location or a part of the child's body.

377 State v. Michaels, 642 A.2d 1372, 1383 (N.J. 1994), *aff'g* State v. Michaels, 625 A.2d 489, 515 (N.J. Super. Ct. App. Div. 1993) ("The record of available interviews does not disclose that any of the children related their testimony of the alleged abuse by 'free recall.'"). The appeals court in *Michaels* acknowledged the conclusion of some experts that the "'development of accurate recall skills' does not occur until the child is five years of age." 625 A.2d at 516 (quoting State v. Wright, 775 P.2d 1224, 1227 (Idaho 1989), *aff'd,* 497 U.S. 805 (1990)). The court failed, however, as did the Idaho court that it cited, to recognize that if a child lacks accurate recall skills, leading questions are a virtual necessity.

strates the reliability of the child's testimony by clear and convincing evidence.[378]

V
THE SCIENTIFIC STANCE OF THE NEW WAVE

Although expert defense witnesses always have presented a challenge to the successful prosecution of child abuse, prosecutors have found it easy to impeach these experts by suggesting that their views on the suggestibility of children reflect value judgments rather than impartial scientific opinion. Ironically, the new wave's scientific stance increases its potential for use as a powerful weapon for defense attorneys. One cannot easily dismiss the new wave's research as the work of defense-oriented zealots. Not only does the new wave conduct research with scientific rigor, but it also admirably discusses its research and its implications in an even-handed tone. Ceci and Bruck have positioned themselves as centrists in a debate between extremes.[379]

To illustrate the comparative credibility of the new wave, consider two of the most prominent veteran expert witnesses in child abuse cases: Ralph Underwager and Richard Gardner. The National Center for Prosecution of Child Abuse named Ralph Underwager enemy expert number one in 1986.[380] He is the director of the Institute for Psychological Therapies and has co-authored numerous books and articles, including three on allegations of sexual abuse with his colleague Hollida Wakefield.[381] Gardner has "conducted consultations and provided testimony in a dozen states and . . . [has] lectured to legal and mental health professionals in 20 more."[382] He is a clinical professor of child psychiatry at the College of Physicians and Surgeons at Columbia University and is the author of over 250 books and articles on child psychotherapy. Specifically, he has written three books on child sexual abuse allegations.

Unlike Ceci and others in the new wave, Underwager and Gardner are not researchers by profession. Underwager and his coauthor,

---

378 *See Michaels*, 642 A.2d at 1383.

379 *See* CECI & BRUCK, *supra* note 36, at 4 ("[W]e make the case that the needs of both science and society dictate a middle ground."); Ceci & Bruck, *supra* note 9, at 433 ("Extreme statements that some have profered [sic] in the media . . . are not supported by the findings reviewed here."); Ceci et al., *supra* note 33, at 494 (noting that "[a] major goal [of] the amicus brief in *State v. Michaels* . . . was to . . . argue that a middle ground can and should be the framework for future investigations").

380 *See* HOLLIDA WAKEFIELD & RALPH UNDERWAGER, RETURN OF THE FURIES: AN INVESTIGATION INTO RECOVERED MEMORY THERAPY 5 (1994).

381 *See, e.g.*, HOLLIDA WAKEFIELD & RALPH UNDERWAGER, ACCUSATIONS OF CHILD SEXUAL ABUSE (1988); WAKEFIELD & UNDERWAGER, *supra* note 380.

382 RICHARD A. GARDNER, SEX ABUSE HYSTERIA: SALEM WITCH TRIALS REVISITED 2 (1991).

Wakefield, pay close attention to the research literature and declare their adherence to the "rational, critical scientific mode of thought that seeks ever closer approximations of truth."[383] However, they remain relatively unimpressed with the recent wave of research documenting the suggestibility of children because "[t]his fact . . . has, of course, been perfectly familiar to those experienced in raising children, for thousands of years."[384] Underwager's most notable contribution to the research on child sexual abuse and child interviewing was a study on anatomical dolls, which Underwager and his colleagues published in their journal, *Issues in Child Abuse Accusations*.[385] The study has received some criticism by the research community.[386]

Gardner is even less impressed than Underwager with research. In his book, *Sex Abuse Hysteria: Salem Witch Trials Revisited*, Gardner excuses his lack of supporting references to research by noting that he easily could find references going either way and that "[t]he term scientific proof is not applicable to most of the issues discussed" in his book.[387] Like many seasoned clinicians, Gardner is quick to rely on his extensive experience with children to overcome any contradictions in the research literature. For example, commenting on Goodman's claim that young children in her research are resistant to sexual questions, Gardner notes "[t]his has not been my experience."[388]

Underwager and Gardner do not hide their value judgments regarding the relative weights one ought to give to false allegations and false denials. Building on Blackstone's remark that it is better to let ten guilty men go free than to let one guilty man go to jail, Gardner argues that "'it is better to let 100 guilty men go free than to convict one innocent man.'"[389] Even more remarkably, Underwager declares that "[i]t is more desirable that a thousand children in abuse situations are not discovered than it is for one innocent person to be con-

---

383 WAKEFIELD & UNDERWAGER, *supra* note 381, at xxi.

384 WAKEFIELD & UNDERWAGER, *supra* note 380, at 34.

385 *See* William McIver II et al., *Behavior of Abused and Non-Abused Children in Interviews with Anatomically Correct Dolls*, ISSUES CHILD ABUSE ACCUSATIONS, Winter 1989, at 39.

386 *See, e.g.*, CECI & BRUCK, *supra* note 36, at 166 ("It is important to point out that this last report has been criticized for the failure to differentiate between explicit sexual behavior and aggressive behavior."); Mark D. Everson & Barbara W. Boat, *Putting the Anatomical Doll Controversy in Perspective: An Examination of the Major Uses and Criticisms of the Dolls in Child Sexual Abuse Evaluations*, 18 CHILD ABUSE & NEGLECT 113, 124 (1994) ("The study, however, has several limitations, the most serious of which is the fact that the authors combined a variety of disparate behaviors into a single category for their group comparison. . . . [The] findings therefore are virtually uninterpretable.").

387 GARDNER, *supra* note 382, at 2.

388 *Id.* at 105.

389 RICHARD A. GARDNER, THE PARENTAL ALIENATION SYNDROME AND THE DIFFERENTIATION BETWEEN FABRICATED AND GENUINE CHILD SEX ABUSE 175-76 (1987).

victed wrongly."[390] The harms of false allegations are self-evident: innocent people are jailed, and once-intact families are destroyed. Gardner believes there are hundreds (and possibly thousands) of wrongly convicted defendants in jail and that "[t]here are hundreds . . . who have committed suicide because of a false sex-abuse allegation."[391]

Although Gardner states that allegations of sexual abuse in general tend to be true, his work emphasizes the problems of allegations in custody disputes, day-care center cases, and nursery-school cases, the majority of which he believes are false.[392] Wakefield and Underwager, for their part, report that sixty percent of the child sexual abuse cases they had dealt with were false allegations.[393]

The harms resulting from false denials also seem obvious: an abused child remains in the abusive home, while a molester goes undetected and unpunished. Labelling these outcomes as "harms," however, presupposes that sexual abuse is itself detrimental, which is something that Gardner and Underwager are willing to question. Gardner, writing in Underwager's journal, notes that "[s]exual activities between adults and children are a universal phenomenon . . . . [S]uch encounters are *not* necessarily traumatic. The determinant as to whether the experience will be traumatic is the social attitude toward these encounters."[394] Underwager goes even further:

> Paedophiles spend a lot of time and energy defending their choice. I don't think that a paedophile needs to do that. Paedophiles can boldly and courageously affirm what they choose. They can say that what they want is to find the best way to love. I am also a theologian and as a theologian I believe it is God's will that there be closeness and intimacy, unity of the flesh, between people. A paedophile can say: "This closeness is possible for me within the choices that I've made."[395]

The new-wave researchers present a stark contrast to Gardner, Underwager, and other experts more accustomed to testifying in

---

390 Scott Kraft, *Careers, Reputations Damaged: False Molesting Charges Scar Lives of Accused*, L.A. TIMES, Feb. 11, 1985, at 1.

391 GARDNER, *supra* note 389, at xxvii.

392 *See* GARDNER, *supra* note 382, at 3-6.

393 *See* WAKEFIELD & UNDERWAGER, *supra* note 381, at 292.

394 Richard A. Gardner, *Belated Realization of Child Sexual Abuse by an Adult*, ISSUES CHILD ABUSE ACCUSATIONS, Fall 1992, at 177, 191.

395 *Interview: Hollida Wakefield and Ralph Underwager*, 3 PAIDIKA 2, 3-4 (1993). In the interview, Wakefield expresses her disagreement with Underwager: "I guess I do feel differently about some things. For example, I find it difficult to envision how a paedophile relationship can have the potential of being the type of close, intimate, constantly developing relationship that would be possible in more traditional relationships . . . ." *Id.* at 4. In their latest book, Wakefield and Underwager argue that "[e]ven though the data seem to suggest otherwise, we maintain that sexual abuse is always harmful." WAKEFIELD & UNDERWAGER, *supra* note 380, at 63.

court. As scientists, they avoid asserting value preferences and recognize the tradeoff between false positives and false negatives.[396] They argue that when one considers the utility of any particular interviewing strategy, one must compare the percentage of abused children who truthfully will reveal, given such a strategy, with the percentage of nonabused children who falsely will reveal given the same strategy.[397] Because these considerations involve empirical questions, research becomes the tool by which practitioners might choose among different styles of interviewing.

Sometimes, however, Ceci acts as if the research literature provides a sufficient ground upon which to make a judgment regarding interviewing style. He speaks of numerators focusing on the likelihood of false negatives and denominators focusing on the likelihood of false positives:

> I . . . do not make any apologies for being a denominator: my best reading of the corpus of scientific research leads me to worry about the possibility of false allegations. It is not a tribute to one's scientific integrity to walk down the middle of the road; the data are more to one side. As I hope to show, the data are somewhat off-center . . . .[398]

The scientific data, however, cannot tell an interviewer how aggressively to interview a child. In addition to the data concerning the relative risks of false positives and false negatives, one also needs to know the relative harms of each type of error. How harmful is a false denial of abuse compared to a false allegation? If a method of inter-

---

396 *See* Ceci et al., *supra* note 33, at 494 ("The trade-off between doing all that one can do to elicit a report of a potentially important event from a child versus avoiding all elicitation techniques that might contaminate the child's recollection presents a conflict."). Ceci and Bruck also expressed the following view:

> Both of these arguments fail to consider the combined costs and benefits of passive versus aggressive interviewing practices. For example, before either of these claims can be substantiated, it is important to determine the proportion of abused children who are initially too scared or confused to divulge the details of their victimization, but who will eventually do so if they are questioned more aggressively, as well as the proportion of nonabused children who will eventually disclose false details of abuse if they are aggressively questioned.

CECI & BRUCK, *supra* note 36, at 2.

397 *See* CECI & BRUCK, *supra* note 36, at 2.

398 Ceci, *supra* note 32, at 18. As discussed earlier, *see supra* notes 379, 396 and accompanying text, Ceci and his colleagues consider themselves centrists. Who are the extremists? On the side of skepticism, an extremist believes the data are sufficient to "categorically discredit children from testifying or even to recommend skepticism upon hearing a child's disclosure." CECI & BRUCK, *supra* note 36, at 4. By Ceci and Bruck's account, Underwager and Gardner do not qualify. Ceci and Bruck cite them as believing that children are "potentially" less reliable than adults, but also as acknowledging that "children are capable of high levels of accuracy, provided that adults who have access to them do not attempt to bias their reports." Ceci & Bruck, *supra* note 9, at 403 n.1. Apparently, centrists consist of a large and heterogeneous lot.

viewing leads thirty percent of abused children to report their abuse and leads ten percent of nonabused children falsely to assert abuses that never occurred, then the method has some diagnostic utility in distinguishing between abused and nonabused children. Such diagnostic utility stems from the fact that a higher percentage of abused children than nonabused children would allege abuse given the method of interviewing. However, the method may produce a large number of false allegations. For example, assume that we interview one hundred children, of whom fifty suffered abuse and fifty did not. Our interviewing method would elicit twenty allegations of abuse, of which five would be false. Is a twenty-five percent rate of false allegations acceptable? Blackstone, Underwager, and Gardner all would say no.

Ceci and his colleagues argue that "[i]t is for the court or the fact finder, not researchers, to decide how much value to attach to each type of error. Researchers who assume such a role are usurping the judicial process."[399] The new wave never underestimates the harms of either false denial or false allegation. As "both scientists and parents," Ceci and Bruck are acutely aware of the need to protect children from the trauma of sexual abuse.[400] On the other hand, they are equally aware of the devastating effects of a false allegation on the falsely accused, even if the allegation does not lead to a criminal conviction.[401] Ceci and Bruck consider themselves "numerator *and* denominator watchers."[402]

Reflecting the available data, the new wave consistently speaks of some "possibility" of false allegations, rather than asserting that false allegations are probable or even frequent. Whereas Gardner and Underwager are willing to assert numbers, Ceci and Bruck prudently avoid such specificity because "reliable data on the frequency of false claims of sexual abuse" do not exist.[403] Ceci and Bruck are willing to assert that "false claims exist, and perhaps in nontrivial numbers,"[404] but "[they] have never claimed either explicitly or implicitly that children's allegations of sexual abuse are often false."[405] Indeed, Ceci recently stated in an interview: "Not only do I believe children can be

399 Ceci et al., *supra* note 33, at 504.

400 CECI & BRUCK, *supra* note 36, at 4.

401 *See* Ceci & Bruck, *supra* note 9, at 421 (listing potential outcomes of unjust prosecution based on false allegation).

402 CECI & BRUCK, *supra* note 36, at 4.

403 *Id.*

404 *Id.*

405 Ceci et al., *supra* note 33, at 499. Appearing on *20/20*, Bruck agreed with the reporter that "there are dozens of people in jail now who are totally innocent." *20/20: From the Mouths of Babes*, *supra* note 46. She subsequently bas explained that the statement was merely ber "personal opinion," that it was taken out of context, and that she was merely drawing the conclusion based on the overall number of prisoners who are innocent and

reliable in sexual abuse cases, I believe the vast majority of them are reliable in those cases."[406]

What then can the new wave tell us about interviewing? Should interviewers be less aggressive than they have been in the past because of the number of false allegations, or should they be more aggressive because so many true claims of abuse remain unreported? Whether as scientists, parents, or both, Ceci and Bruck offer recommendations for interviewers, despite the implicit value judgments that these recommendations entail. For example, they make the facially unexceptional recommendation that "the ideal interview should not contain techniques that have been found to have harmful consequences."[407] To justify their recommendations from a scientific stance, Ceci and Bruck invoke a powerful analogy between interviewing methods and a cure for cancer.[408] If everyone's moral senses respond in the same fashion to the analogy, then the authors can claim that they are not imposing their own value judgments on others by making recommendations for interviewing practice.

The analogy is as follows: Suppose there exists a drug that cures some with cancer, but causes cancer in some who would otherwise be cancer-free. Would one give the drug to everyone, assuming one has no idea who has cancer and who does not? "Obviously not."[409] The drug is analogous to suggestive techniques in interviewing. One defends these techniques on the grounds that they increase the likelihood that abused children will reveal their abuse. The techique, however, also may create false allegations. If the analogy holds, most people would agree that interviews should not use these techniques.

On the one hand, the analogy is evenhanded. Note that the harm of a false positive and of a false negative is the same—cancer. On the other hand, the analogy omits exactly what we hope scientists will provide—the percentage of those with cancer whom the drug will cure and the percentage of those without cancer who will develop the disease. Without these percentages, the answer to the question of whether we should provide the drug to everyone is not so obvious. The answer is obvious only if we assume that we should not give a drug to individuals when it can cause the harm we are seeking to avoid. In terms of interviewing children, the analogy would imply that *any* false allegation of sexual abuse produced by an interviewing method ren-

---

the overall number of prisoners serving time for sexual crimes. Bruck, *supra* note 44, at 100.

406 *Nightline*, *supra* note 133.

407 Bruck et al., *supra* note 30, at 148.

408 *See* Ceci, *supra* note 32, at 16-17.

409 *Id.*

ders that method unacceptable, no matter how many true allegations the method elicits.

Another objection to the analogy stems from the unjustified assumption that we would administer the drug to everyone, even those who showed no signs or symptoms of cancer. In terms of sexual abuse, this assumption would mean that we interview children about whom we have no suspicion of abuse. In fact, investigative interviewers usually question only those children whom they suspect have been abused. An important, and often overlooked, fact is that as the proportion of truly abused children among those researchers interview varies, so does the number of false allegations any interviewing method produces.[410] Recall the example in which we interviewed 100 children with a method that elicited reports of abuse in thirty percent of the abused children and ten percent of the nonabused children. When we assumed that *half* of the children in fact had suffered abuse, we obtained twenty allegations of abuse, twenty-five percent of which were false. If we assume instead that *eighty percent* of the children interviewed in fact had suffered abuse, then our method would elicit twenty-four true reports (thirty percent of the eighty abused children) and two false reports (ten percent of the twenty nonabused children). Now only eight percent of the allegations we elicit would be false. As the number of truly abused children among those that researchers interview grows, the percentage of false allegations decreases.

In deciding whether an interviewing method is unduly suggestive, we therefore must speculate about the likely ratio of abused to nonabused children in the pool of interviewed children. Anticipating this problem, Ceci offers a hypothetical in which we are eighty percent certain that those to whom we give the drug have cancer.[411] This hypothetical analogizes to the situation in which eighty percent of the children interviewed in fact have suffered abuse. Without positing anything about the relative rates at which abused and nonabused children would disclose abuse, Ceci argues that we would "[p]robably not" administer the drug.[412] Therefore, lowering the percentage of allegations that are false by narrowing the class of interviewed children does not change Ceci's value judgment that the costs of potential

---

410 *Cf.* James M. Wood, *Weighing Evidence in Sexual Abuse Evaluations: An Introduction to Bayes's Theorem*, 1 CHILD MALTREATMENT 25, 27 (1996) ("[I]t is not enough to know that evidence is weak or strong. Exactly the same evidence may lead to quite different conclusions, depending on the rate of abuse in the group being evaluated.").

411 *See supra* text accompanying note 409.

412 Ceci, *supra* note 32, at 17. *But cf.* CECI & BRUCK, *supra* note 36, at 3 ("Is the diagnostic test for determining the 80% risk status valid? And, if so, is the risk of infecting the 20% . . . a price we are willing to pay for the chance to treat the 80% who do? These are questions about which reasonable people can and do disagree.").

false allegations are not worth the benefits of revealing some cases of actual child abuse.

In their book, Ceci and Bruck offer a slightly modified version of the same analogy.[413] Rather than imagine a drug that cured cancer in "some" and caused cancer in "some," the authors hypothesized a drug that "prevented" cancer by curing everyone with cancer but "created" cancer in everyone who was cancer-free, apparently assuming that the drug cures everyone with cancer and causes cancer in everyone without cancer. "Assuming no reliable method exists for detecting which individuals have cancer and which do not, should the drug be administered to everyone? Probably not . . . ."[414] In this version of the analogy, Ceci and Bruck have specified the relative rate at which false positives and false negatives occur, but these rates are unrealistic. The analogy only would apply to a procedure by which interviewers classified all children as abused, so that all abused children would be correctly classified and all nonabused children would be incorrectly classified. This procedure would be completely nondiagnostic as a means of determining which children have suffered abuse.

The new wave emphasizes the possibility, rather than the probability, of false allegations when converting research results into policy recommendations. Ironically, by refusing to argue that false allegations are common, and by refusing to quantify the proportion of acceptable false allegations, the new wave goes beyond Underwager and Gardner, while at the same time adopting a more rational tone. If the mere existence of false allegations justifies reducing the number of true allegations we elicit, then Ceci and Bruck's version of Blackstone's ratio is very high indeed.

An emphasis on the *possibility* (rather than relative *probability*) of false allegations has other advantages. Much of the debate over the new wave of suggestibility research is based on the extent to which one can translate the research to the real world. Are the methods experimenters use in the research the same as those most interviewers use in actual child abuse investigations? Would the children respond to the experimenters' questions the same if they were asked about sexual abuse? These questions become unimportant if we focus on the *possibility,* rather than the *probability,* of false allegations. One need not prove that most interviewers use coercive methods, only that some do. One need not prove that children who report nonevents in the new wave's research equally will allege abuse falsely, only that some will.

In the legal arena, one might argue that an emphasis on possibility is justifiable because possibility constitutes reasonable doubt. If

413 *See* CECI & BRUCK, *supra* note 36, at 3.

414 *Id.*

this argument held true, however, then juries always should acquit because wrongful conviction is always a possibility. Prosecutors never conclusively prove guilt, but guilty verdicts are common. Therefore, jurors must consider the need to minimize false acquittals as well as false convictions. Civil cases illustrate this tradeoff even more clearly because they tolerate a much higher likelihood of false positive error. One cannot quantify easily how large an error is tolerable; some even argue that we should not attempt to do so.[415] For the judicial system to function, the tolerable risk of error must be greater than zero.

Even if one decides to tolerate false convictions, a single-minded focus on the possibility of false positives would distort the process by which courts make judgments. The new-wave research is likely to have the greatest impact on judges' decisions regarding the admissibility of evidence: whether they should allow children to testify, whether they should admit children's out-of-court statements, and whether they should allow experts to testify regarding suggestibility. In each situation, judges weigh the risks of various types of error in deciding whether to admit the proffered evidence, but do not need to decide whether abuse in fact occurred. Rather, judges will admit evidence as long as it has some probative value (i.e., it increases the likelihood that abuse occurred), and its probative value is not substantially outweighed by its prejudicial effect (i.e., the jury will not give it more weight than it actually deserves).[416]

Assessing the probative value of evidence requires some understanding of both false positives and false negatives. Claims of abuse have some probative value as long as abused children are more likely to claim abuse than nonabused children.[417] The possibility of false allegations merely means that some nonabused children will claim abuse, not that claims of abuse are irrelevant.

Assessing prejudice requires the court to consider whether jurors are likely to give certain types of evidence too much weight. An emphasis on the possibility of false allegations would be justified if jurors are prone to accept uncritically claims of abuse as true—if they as-

---

415 *See* Laurence H. Tribe, *Trial by Mathematics: Precision and Ritual in the Legal Process*, 84 HARV. L. REV. 1329, 1331-32 (1971).

416 *See* FED. R. EVID. 401, 403 (providing the standards for relevance and undue prejudice). Hearsay faces heightened standards of admissibility because of its potentially prejudicial effects. When assessing the admissibility of a child's out-of-court statement under various hearsay exceptions, a court considers whether factors exist that compensate for the opponent's inability to test the truth of the statement by the trial process (particularly cross-examination). Essentially, the concern is that the jury may give the untested statement more weight than it deserves, making it unduly prejudicial.

417 This statement is logically equivalent to the definition of "[r]elevant evidence." FED. R. EVID. 401; *see* Thomas D. Lyon & Jonathan J. Koehler, *The Relevance Ratio: Evaluating the Probative Value of Expert Testimony in Child Sexual Abuse Cases*, 82 CORNELL L. REV. 43, 46 (1996).

sume that there is no such thing as a false positive. Ceci and Bruck justify their focus on the "weaknesses" of children's testimony on the grounds that the weaknesses "are less well understood by experts and nonexperts" than the strengths.[418] Various surveys of jury-eligible citizens, however, demonstrate that jurors are well aware of the potential for suggestibility.[419] Potential jurors with personal experience of sexual abuse, who may be inclined to believe young children's abuse allegations, are less likely to make it onto the jury.[420] The potential for prejudice from admitting into evidence children's statements appears less serious than that from admitting expert testimony that children are suggestible—something jurors likely already know.

The most troubling aspect of the emphasis on false allegations is that it may overwhelm our awareness of the need to uncover true allegations of abuse. One cannot easily keep both false positives and false negatives in mind at once. Moreover, despite a warning that *possibilities* are not *probabilities*, people tend to estimate the likelihood of events based upon the ease with which they can imagine the events.[421] Making false allegations easily imaginable increases their perceived probability. This increase in perceived probability is especially likely

---

418 CECI & BRUCK, *supra* note 36, at x.

419 *See* Michael R. Leippe & Ann Romanczyk, *Children on the Witness Stand: A Communication/Persuasion Analysis of Jurors' Reactions to Child Witnesses*, *in* CHILDREN'S EYEWITNESS MEMORY, *supra* note 19, at 155, 159 (presenting a survey of parents and college students that found that "the majority of respondents saw 5- to 9-year-old children as more suggestible than adults when the influence agent is an adult"); David F. Ross et al., *Age Stereotypes, Communication Modality, and Mock Jurors' Perceptions of the Child Witness*, *in* PERSPECTIVES ON CHILDREN'S TESTIMONY, *supra* note 221, at 37, 38 (presenting a survey of college students who "believed the child witness, whether six or eight years old, was less likely to be accurate and more likely to be open to suggestion than witnesses of adult age (either young or old)"); A. Daniel Yarmey & Hazel P. Tressillain Jones, *Is the Psychology of Eyewitness Identification a Matter of Common Sense?*, *in* EVALUATING WITNESS EVIDENCE: RECENT PSYCHOLOGICAL RESEARCH AND NEW PERSPECTIVES 13, 33 & tbl.2.15 (Sally M.A. Lloyd-Bostock & Brian R. Clifford eds., 1983) (reporting that laypersons are likely to believe that eight-year-old child is highly suggestible). Review of the research literature on mock jurors' reactions to child witnesses is beyond the scope of this paper. Suffice it to say that the results are mixed and of limited utility in assessing jurors' attitudes about suggestibility; factors other than suggestibility heavily influence their evaluations of child witnesses' credibility, such as their judgments regarding children's proclivity to lie.

420 *See* PAUL STERN, PREPARING AND PRESENTING EXPERT TESTIMONY IN CHILD ABUSE LITIGATION 3 (1997) ("Those who report past victimization [of child sexual abuse] are unlikely to be seated on the jury.").

421 *See* Slovic et al., *supra* note 149, at 467 (reviewing research in which subjects estimated the frequency of various causes of death and concluded that "overestimated causes of death were dramatic and sensational, whereas underestimated causes tended to be unspectacular events, which claim one victim at a time and are common in nonfatal form"); Amos Tversky & Daniel Kahneman, *Availability: A Heuristic for Judging Frequency and Probability*, *in* JUDGMENT UNDER UNCERTAINTY: HEURISTICS AND BIASES, *supra* note 149, at 163, 164 ("A person is said to employ the availability heuristic whenever he estimates frequency or probability by the ease with which instances or associations could be brought to mind.").

when vivid portrayals by videotapes or anecdotes enhance their imaginability. The new-wave researchers are not only scientists, but also storytellers who disseminate their most impressive subjects as aggressively as their cumulative data.[422] When possibilities become probabilities, fears turn into unwarranted legal presumptions that question the reliability of the evidence children present.

CONCLUSION

Great care is needed in how researchers describe their results to minimize the chance that they will be misused by those involved in the adversarial process.
—Stephen J. Ceci & Helene Hembrooke[423]

The goal of this paper is to describe how the scientific stance of the new wave of suggestibility research conceals questionable empirical assumptions and subjective value judgments. Although the new wave asserts that one can generalize its findings to actual abuse allegations, the new wave's research ignores the realities of sexual abuse and of actual abuse investigation. These shortcomings render the new-wave research of limited applicability to real-world abuse cases.

While claiming to address the problems of both false allegations and false denials of abuse, the new wave emphasizes the dangers of false allegations and advocates changes in interviewing strategies that may make it more difficult to detect true cases of abuse. My hope is that the courts and others will carefully scrutinize the claims of these social scientists regarding the suggestibility of young children. My position is not that the new wave's research is irrelevant to decision makers assessing the reliability of children's claims of abuse, but that its relevance is tempered by the realities of sexual abuse and abuse investigations and by the fact that no science is value-free.

One could say much more about the new wave and the implications of its research for sexual abuse cases. For example, I have not discussed the controversies surrounding repeated interviews and the use of anatomically correct dolls. I also have not addressed the significance of age differences in suggestibility. Careful consideration of these issues might make real-world interviews look worse and the new-wave research look better. Alternatively, this consideration might reveal other difficulties in applying the result of this research to the real

422 *See* Bruck & Ceci, *supra* note 39, at 294 (discussing pilot subject whose behavior "demonstrates vividly the potential suggestiveness of anatomical dolls with non-abused 3-year-olds" in an amicus brief submitted to the New Jersey Supreme Court, and appending to the brief a videotape on the subject's behavior).

423 Stephen J. Ceci & Helene Hembrooke, *The Contextual Nature of Earliest Memories, in* MECHANISMS OF EVERYDAY COGNITION 117, 124 (James M. Puckett & Hayne W. Reese eds., 1993).

world.[424] Resolution of these issues requires further critical analysis. Further discussion is especially timely because of the growing attention courts are paying to this research. Generally, courts have been receptive to the new-wave research, leading the new wave to claim that their work is "beginning to have some impact on the legal system in terms of the decisions that are made by trial and appellate courts."[425] Reviewing the two cases that the new wave cites to support this claim confirms that the courts are skeptical of the veracity of young children's claims of abuse after extensive interviewing has taken place.

In *State v. Michaels*,[426] Ceci and Bruck authored an amicus brief (cosigned by forty-three well-respected social scientists) on behalf of the Committee of Concerned Social Scientists. The brief emphasized the suggestiveness of pretrial interviews in light of recent research on suggestibility, much of it conducted by the new wave. Psychologists have noted that the New Jersey Supreme Court's opinion "frequently referred" to the amicus brief,[427] and two legal scholars have asserted that the "brief obviously educated the supreme court of New Jersey and, in so doing, helped bring the legally sanctified torture of Ms. Michaels to an end."[428] The court held that if a criminal defendant can demonstrate a substantial likelihood that a child witness's testimony was the product of pretrial suggestion, the child cannot testify unless the state establishes the testimony's reliability by clear and convincing evidence.[429] The case constitutes an unprecedented limitation on child witnesses' testimony.

In *United States v. Rouse*,[430] the Eighth Circuit Court of Appeals held that preventing an expert psychological witness from testifying that pretrial interviewing techniques constituted a "practice of suggestibility" was reversible error.[431] The opinion frequently referred to Ceci and Bruck's 1995 book and their 1993 review article, including a

---

424 *See* Thomas D. Lyon, Questioning Children: The Effects of Suggestive and Repeated Questioning (1998) (unpublished manuscript, on file with author) (discussing the effects of repeating questions within and between interviews); Karen J. Saywitz & Thomas D. Lyon, Coming to Grips with Children's Suggestibility: Confronting Limitations, Promoting Capabilities (1998) (unpublished manuscript, on file with author) (discussing the importance of age differences in suggestibility).

425 Bruck et al., *supra* note 30, at 137 (citing United States v. Rouse, 100 F.3d 560 (8th Cir. 1996), *aff'd en banc*, 111 F.3d 561 (8th Cir. 1997); State v. Michaels, 642 A.2d 1372 (N.J. 1994)).

426 642 A.2d 1372 (N.J. 1994).

427 Warren et al., *supra* note 153, at 231; *see also* CECI & BRUCK, *supra* note 36, at 337; C.J. Brainerd & D. Hill, *Voices of Children*, 42 CONTEMP. PSYCHOL. 7 (1997) (reviewing CECI & BRUCK, *supra* note 36) (stating that the brief "was quoted by the court in its decision").

428 Ronald J. Allen & Joseph S. Miller, *The Expert as Educator: Enhancing the Rationality of Verdicts in Child Sex Abuse Prosecutions*, 1 PSYCHOL. PUB. POL'Y & L. 323, 336 (1995).

429 *See Michaels*, 642 A.2d at 1384-85.

430 100 F.3d 560 (8th Cir. 1996), *aff'd en banc*, 111 F.3d 561 (8th Cir. 1997).

431 *Id.* at 569.

reference to the book's discussion of the Salem witch trials and its implications for modern day discussions of children's credibility.[432] Ceci and Bruck's reviews of the research enabled the court to conclude that the expert's proffered testimony fulfilled the requirements of *Daubert v. Merrell-Dow Pharmaceuticals*[433] that expert scientific testimony be both reliable and helpful to the trier of fact[434] because the suggestive methods that the expert would discuss had "been amply demonstrated in the psychological literature as producing undue suggestibility in children's testimony."[435]

Closer analysis of these cases, however, suggests that these courts had an ambivalent reaction to the new wave of suggestibility research. Although the amicus brief Bruck and Ceci submitted to the New Jersey Supreme Court may have influenced the court's opinion, it is remarkable that the opinion *never* cited the brief, notwithstanding the claims of some commentators.[436] The New Jersey opinion also failed to mention any of the new-wave research this Article discusses, including the Sam Stone Study, the Mousetrap Study, and the Inoculation Study.[437] Rather, the social scientist the court cited most frequently was Gail Goodman, one of the few researchers who refused to sign the amicus brief.

The omission likely was not an oversight. The New Jersey Supreme Court was at pains to avoid taking sides in the debate over children's suggestibility. The court emphasized a "fairly wide consensus" among "experts, scholars, and practitioners concerning improper interrogation techniques in finding that the interviewing practices at issue in *Michaels* were improper."[438] By repeatedly quoting Goodman, whom the appellate court had characterized as making recommendations "slanted in favor of [the] prosecution of sex abuse cases,"[439] the court clearly hoped to render its conclusions uncontroversial. Furthermore, by failing to quote the amicus brief, the court implicitly recognized that the "concerned social scientists" were interested in policy as well as in science.

The new wave's apparent success in *Rouse* is particularly equivocal. The testifying expert in *Rouse* was not Ceci, Bruck, or any of the experimental psychologists who signed the amicus brief in *Michaels.*

---

432 *See id.* at 563 n.2.

433 509 U.S. 579 (1993).

434 *See Rouse,* 100 F.3d at 567-73.

435 *Id.* at 569.

436 *See* sources cited *supra* note 427.

437 The case cites Ceci once, for the proposition that younger children are more suggestible than older children. *See* State v. Michaels, 642 A.2d 1372, 1378 (N.J. 1994).

438 *Id.* at 1376.

439 State v. Michaels, 625 A.2d 489, 511 (N.J. Super. Ct. App. Div. 1993), *aff'd,* 642 A.2d 1372 (N.J. 1994).

Rather, it was Ralph Underwager, the psychologist who has publicly proclaimed that "[i]t is more desirable that a thousand children in abuse situations are not discovered than it is for one innocent person to be convicted wrongly"[440] and that "[p]aedophiles can boldly and courageously affirm what they choose."[441] Second, the appellate court did not question the district court's ruling that Underwager could only "express his own expert opinions and explain his own prior research" on the grounds that "there is not anywhere near yet the agreement in the [scientific] community as to methods, techniques, testing or reliability that would warrant the admissibility before a jury of these matters."[442] Third, the original opinion in *Rouse* was vacated, a rehearing granted, and the en banc court held on rehearing that limiting Underwager's testimony was harmless error.[443] As a result, the endorsement of Ceci and Bruck's reviews is now in a dissenting opinion.[444]

The new wave's most recent victory concerns Cheryl Amirault LeFave's conviction in the Fells Acres' daycare case. Relying largely on the affidavit and testimony of Bruck, a Massachusetts Superior Court judge overturned LeFave's conviction, holding that recent research on children's suggestibility constitutes "new evidence" proving that suggestive interviewing practices "forever tainted" the testimony of the child witnesses.[445] The interviews the court quoted in the opinion read like those in *Michaels*—highly suggestive, even coercive, questions that evince a single-minded determination to uncover abuse at all costs. Assuming that the appellate courts agree with the lower court that the interviewing practices deprived LeFave of due process, the difficulty they face is how to do justice in the case before them without doing injustice to the thousands of garden-variety sexual abuse cases that are prosecuted with little fanfare. What legal rule will protect the innocent without freeing the guilty? Striking a balance requires an empirical judgment regarding the suggestibility of children and a value judgment regarding the tradeoff between false convictions and false acquittals. Those are judgments about which reasonable people—even scientists—long will disagree.

---

440 Kraft, *supra* note 390, at 14.

441 *Interview: Hollida Wakefield and Ralph Underwager*, *supra* note 395, at 4.

442 United States v. Rouse, 100 F.3d 560, 583 (8th Cir. 1996) (Loken, J., dissenting) (quoting the district court opinion), *aff'd en banc*, 111 F.3d 561 (8th Cir. 1997).

443 . *See* United States v. Rouse, 111 F.3d 561, 572 (8th Cir. 1997) (en banc).

444 *See id.* at 576 (Bright, J., dissenting).

445 *See* Findings of Fact, Rulings of Law, and Order on Defendant's Motion for a New Trial, Commonwealth v. LeFave, No. 85-63 (Mass. Super. Ct. June 1998).