

University of Southern California Law

From the SelectedWorks of Thomas D. Lyon

January 2002

6. Lyon, T.D. (2002). Applying suggestibility research to the real world: The case of repeated questions. *Law & Contemporary Problems*, 65, 97-126.

Contact
Author

Start Your Own
SelectedWorks

Notify Me
of New Work



Available at: <http://works.bepress.com/thomaslyon/7>

APPLYING SUGGESTIBILITY RESEARCH TO THE REAL WORLD: THE CASE OF REPEATED QUESTIONS

THOMAS D. LYON *

I

INTRODUCTION

One can discern two parallel trends in the law and the psychology of child witnesses. In the law, appellate courts are beginning to stem the once powerful movement to increase the acceptance of children's testimony and the admissibility of children's out-of-court statements. In psychology, experimental psychologists are amassing evidence of the potential unreliability of children's memory reports. The trends intersect when courts assess the reliability of children's statements in order to evaluate the competency of child witnesses, to decide whether to admit expert testimony about the suggestibility of children, and to decide whether to admit children's hearsay.

This article will analyze particular strands of each trend. With respect to the law of the child witness, it will consider the application of the medical diagnosis hearsay exception to sexual abuse cases, using as a case study *State v. Larson*,¹ a Minnesota case that made its way to the United States Supreme Court. The analysis will show how restrictive application of the medical diagnosis exception forces courts to confront the dangers of children's suggestibility.

With respect to the psychology of child witnesses, this article will consider the application of the research literature on repeated questions to sexual abuse cases. It will review the entire corpus of research on repeated questions and apply that research to *State v. Larson*. The article will argue that the risks of question repetition have been exaggerated. The leading research on repeated questions does not support a claim that repetition increases error. Whether repetition leads to inconsistency depends on the types of questions asked, the age of the child, and the child's memory of the event. Most important, researchers ignore the potential effects of repetition on false denials, emphasizing instead the risk that repetition will lead to false allegations. As a result, the

Copyright © 2002 by Thomas D. Lyon

This article is also available at <http://www.law.duke.edu/journals/65LCPLyon>.

* Professor of Law, University of Southern California.

Thanks to Catherine Halvorsen of the USC Law Library for her research assistance, to Michael Cromett, Assistant State Public Defender, State of Minnesota, and to Paul Kempainen, Assistant Attorney General, State of Minnesota, for sharing their briefs in *State v. Larson*. Thanks to Dan Simon for commenting on an early draft.

1. 453 N.W.2d 42 (Minn. 1990).

cost-benefit calculations called for by policy-minded researchers include only the costs of repeated questions.

Part II discusses the facts of *State v. Larson*, the application of the medical diagnosis exception, and how limiting the exception leads courts to look more carefully at the potential unreliability of children's reports. A key issue in *Larson* was whether the child's response to repeated questions could be trusted. Part III discusses the importance of repeated questions in debates over suggestibility and summarizes the settled view of the dangers of repeated questions. Part IV critically reviews the literature, including the factors that affect the likelihood that repetition leads to inconsistency or error. Part V discusses the possibility that repetition will increase the number of true reports of abuse. The goals of this article are to communicate the specifics of the research for practical application in future cases and to illustrate the difficulties and potential drawbacks of applying suggestibility research to real-world cases.

II

STATE V. LARSON: THE MEDICAL DIAGNOSIS EXCEPTION, RELIABILITY, AND REPEATED QUESTIONS

In *State v. Larson*,² Bruce Larson was found guilty of sexually assaulting his four-year-old daughter in the second degree. When the child, hereinafter "B.", was three years old, she complained to her mother of vaginal soreness and burning urination.³ Within a week or so, the mother took the child to a family practice clinic, where the child was examined by a physician's assistant:⁴

The physician's assistant, who conducted the examination, noticed redness and swelling around the vaginal opening. Suspecting abuse, the examiner asked B. who had touched her. B. replied that "Daddy" had touched her with "his bone" that "came out of his pants." B. said it had hurt and that she had asked him to stop.⁵

On appeal, the Minnesota Supreme Court upheld the admissibility of the child's statements to the physician's assistant under the medical diagnosis exception to the hearsay rule.⁶ The court also held that statements the child later made to other adults qualified under the residual exception to the hearsay rule.⁷ The child did not testify at the trial.⁸

The subsequent history of the case illustrates the greater scrutiny courts are giving children's allegations of sexual abuse. The defendant sought review in the U.S. Supreme Court, which granted certiorari and remanded the case,⁹ instructing the Minnesota Supreme Court to reconsider its holding in light of

2. *Id.*

3. *Id.* at 43.

4. *Id.*

5. *Id.*

6. *Id.* at 47.

7. *Id.* at 46-47.

8. *Id.* at 44.

9. *Larson v. Minnesota*, 498 U.S. 801 (1990).

Idaho v. Wright,¹⁰ which held that for statements of a non-testifying declarant to be admissible under the residual exception to the hearsay rule, the court had to find that the statements had “particularized guarantees of trustworthiness,” because the residual exception is not a “firmly-rooted” exception to the hearsay rule.¹¹ In *Wright*, the Supreme Court suggested that trustworthiness could be guaranteed by, among other things, the “lack of [a] motive to fabricate” and by the “spontaneity and consistent repetition” of the statements.¹²

What puzzled the Minnesota Supreme Court on remand was that the most important hearsay—the child’s statements to the physician’s assistant—had been admitted under the medical diagnosis exception rather than under the residual exception.¹³ *Wright* said nothing about how statements admitted under the medical diagnosis exception were to be handled. The Minnesota Supreme Court acknowledged Professor Robert Mosteller’s argument that some statements admitted under the medical diagnosis exception do not fall within the “firmly-rooted” core of the exception, and therefore should be subject to the “particularized guarantees of trustworthiness” test.¹⁴ Although the court believed that the child’s statements to the physician’s assistant fell within the “core” part of the medical diagnosis exception—making an analysis of trustworthiness unnecessary—it proceeded to assess the trustworthiness of the child’s statements.¹⁵

In their briefs on remand, the parties vigorously debated the reliability of the child’s statements. Calling the Minnesota Supreme Court’s recitation of facts “misleading,” the defendant argued that the physician’s assistant admitted in cross-examination that she “asked the child ‘Did your Daddy touch you’ and that the child initially answered ‘no’ and then recanted and said ‘yes’.”¹⁶ By contrast, the State and the court on remand emphasized the physician’s assistant’s direct testimony and her dictated report,¹⁷ which told a different story:

The physician’s assistant, who conducted the examination, saw redness and swelling around the vaginal opening. Suspecting abuse, the examiner, after assuring B. that she was there to “assist her,” asked B. if anyone other than her mother or herself had touched her vaginal area. B. initially responded “no” and turned away. When the ex-

10. *Idaho v. Wright*, 497 U.S. 805 (1990).

11. *Id.* at 816.

12. *Id.* at 821.

13. *State v. Larson*, 472 N.W.2d 120, 126 (Minn. 1991).

14. *Id.* (citing Robert Mosteller, *Child Sexual Abuse and Statements for the Purpose of Medical Diagnosis or Treatment*, 67 N.C. L. Rev. 257, 260-61 (1989) [hereinafter Mosteller, *Child Sexual Abuse*]). The defendant had not challenged whether the use of the medical diagnosis exception was “firmly rooted.” Appellant’s Brief on Remand From the United States Supreme Court at 6, *State v. Larson*, 453 N.W.2d 43 (Minn. 1990) (No. C2-88-2379). Cf. *State v. Jacobsen*, No. 96-8-03132-9, 2000 WL 526749, at *5 (Wash. App. 2000) (“ER 803(a)(4) is not a ‘firmly rooted’ hearsay exception when applied to child hearsay.”); *State v. Florczak*, 882 P.2d 199, 207 (Wash. App. 1994) (“[I]n the context of child hearsay, ER 803(a)(4) is not a firmly rooted hearsay exception.”).

15. *Larson*, 472 N.W.2d at 126.

16. Appellant’s Brief on Remand From the United States Supreme Court at 24-25 n.2, *State v. Larson*, 453 N.W.2d 42 (Minn. 1990) (No. C2-88-2379).

17. Respondent’s Brief on Remand From the United States Supreme Court at 6-7, *State v. Larson*, 453 N.W.2d 42 (Minn. 1990) (No. C2-88-2379).

aminer repeated the question, B. said "yes," that "Daddy" had touched her there. When the examiner asked whether her daddy had touched her there to help wipe her or dry her, B. said "yes," but then, when asked "where," pointed to her abdomen. When asked with what he had touched her, B. replied that "Daddy" had touched her "between [the] legs" with "his bone" that "came out of his pants." Asked whether this had happened before, B. said "yes," that "he laid me on the couch and he rubbed my belly. . . until my skin came off." B. further stated that it hurt "because he pressed hard against her tummy and hugged her real hard," and that she had asked him to stop. She further described feeling wet on her abdomen and having to wipe herself dry afterwards. Although B. apparently told her mother when the examiner briefly left the room that she was "only kidding," B. immediately recanted this to the examiner and explained that "she didn't want to have her father in trouble." Indeed, the examiner observed that B. was "very nervous and very protective of her father," whom B. "obviously loved."¹⁸

The court concluded that the statements enjoyed the "particularized guarantees of trustworthiness," in part because "the questions the assistant asked were not the kind of questions that would prompt a child to fabricate the response the questions elicited."¹⁹ The more complete treatment of the statements given by the court on remand, however, raises several questions about the child's credibility. The child only acknowledged touching when the question was repeated. Although the parties disagree, the question may have been as leading as "Did your Daddy touch you?" The child's subsequent statements were similarly inconsistent, and she was reluctant to name her father as the perpetrator.²⁰ Nevertheless, the court assumed that repetition of the questions was necessary to elicit a truthful allegation from the child.²¹

Under the traditional analysis for the medical diagnosis exception, the court need not have analyzed the trustworthiness of the statements. The lesson of *Larson* is that if a court is forced to do so, the analysis will raise questions about children's suggestibility.²² In some jurisdictions, the medical diagnosis exception contains an explicit requirement that the court assess the reliability of the statements before admitting them.²³ Courts can also inject reliability questions by carefully scrutinizing the motives of the child who made the statement. Assessing the child's motives allows a court to consider whether the statements were the product of leading or suggestive questioning, rather than a desire to receive treatment.²⁴

In many jurisdictions, the statements might fail to qualify as having been made for the purposes of advancing diagnosis or treatment. One can question

18. *Larson*, 472 N.W. 2d at 122.

19. *Id.* at 126.

20. *See id.* at 127.

21. *See id.* at 126.

22. *See id.*

23. *See, e.g., State v. Kay*, 927 P.2d 897, 908 (Idaho Ct. App. 1996) (in applying medical diagnosis exception to child's statements, court can consider "factors that otherwise indicate the reliability of the statements"); MISS. R. EVID. 803(4) (court must find that the "statements were made under circumstances substantially indicating their trustworthiness"); N.H. R. EVID. 803(4) (same).

24. *See State v. Dever*, 64 Ohio St. 3d 401, 410 (1992) (if the statements were "inappropriately influenced by another, then those statements would not have been made for the purpose of diagnosis or treatment.").

whether a four-year-old child understands the importance of telling the truth to doctors.²⁵ Some courts have begun to require an affirmative showing that the child understood the significance of her statements,²⁶ rather than assuming such understanding.²⁷ Especially problematic is the identification of the perpetrator.²⁸ Under the medical diagnosis exception, statements as to the cause of an injury must be reasonably pertinent to diagnosis or treatment.²⁹ Although some courts have held that the identity of the perpetrator is relevant in determining the best course of psychological treatment,³⁰ and in determining whether the child should be removed from her home,³¹ other courts have expressed skepticism that these remedies really constitute “medical” treatment and are covered by the exception.³² Whether identification of the perpetrator is medically relevant from the child’s perspective can also be questioned.³³

25. Indeed, the defendant argued on appeal in *Larson* that a four-year-old child “probably does not understand the importance or need to speak truthfully with a physician which is the fundamental principle behind this exception.” Appellant’s Brief on Remand From the United States Supreme Court at 25 n.2, *State v. Larson*, 453 N.W.2d 42 (Minn. 1990) (No. C2-88-2379). See Mosteller, *Child Sexual Abuse*, *supra* note 14, at 293 (“The age and mental maturity of the child may attenuate the selfish interest of the declarant so profoundly as to virtually eliminate any trustworthiness guarantee under the rationale of this exception.”); 2 JOHN E.B. MYERS, EVIDENCE IN CHILD ABUSE AND NEGLECT CASES § 7.41 at 294 (3d ed. 1997) (“When the [diagnosis or treatment] exception is applied to young children, the developmental fit may not be very good, raising legitimate questions about young children’s understanding of the need for truthfulness with the doctor, and casting doubt on the reliability of certain statements by young children.”). Recent research has suggested that young children are actually quite good at recognizing the role of medical professionals in providing treatment, and, contrary to common belief, do not believe that illness and treatment are forms of punishment. See Melody R. Herbst, Margaret S. Steward, John E.B. Myers & Robin L. Hansen, *Young Children’s Understanding of the Physician’s Role and the Medical Hearsay Exception*, in CHILDREN’S UNDERSTANDING OF BIOLOGY AND HEALTH 235 (Michael Siegal & Candida C. Peterson eds., 1999); Pamela M. Kato, Thomas D. Lyon & Christina Rasco, *Reasoning About Moral Aspects of Illness and Treatment by Preschoolers Who are Healthy or Who Have Chronic Illness*, 19 DEVELOPMENTAL & BEHAV. PEDIATRICS 68 (1998).

26. See, e.g., *State v. Hinnant*, 523 S.E.2d 663, 669 (N.C. 2000) (“[T]he proponent of Rule 803(4) testimony must affirmatively establish that the declarant . . . made the statements understanding that they would lead to medical diagnosis or treatment.”); *United States v. Sumner*, 204 F.3d 1182, 1185 (8th Cir. 2000) (“[I]t must be shown that the child understands the ‘medical significance of being truthful.’”).

27. See, e.g., *United States v. Pacheco*, 154 F.3d 1236, 1241 (10th Cir. 1997) (“[T]he defendant had failed to point to any actual evidence to show the victim did not understand she was seeking medical treatment.”).

28. Surveying the case law, Robert Mosteller argues that, “if what is required is consensus rather than majority view, that consensus is that, in criminal cases, statements identifying the abuser do not fit within” the exception. Robert Mosteller, *The Maturation and Disintegration of the Hearsay Exception for Statements for Medical Examination in Child Sexual Abuse Cases*, 65 LAW & CONTEMP. PROBS. 47 [hereinafter Mosteller, *Maturation*]. Cf. MYERS, *supra* note 25, § 7.39, at 279 (“A small minority of courts reject statements identifying the perpetrator.”).

29. FED. R. EVID. 803(4). According to the advisory committee notes, statements as to fault are usually not pertinent. FED. R. EVID. 803(4) advisory committee’s note.

30. See, e.g., *People v. Falaster*, 670 N.E.2d 624, 629 (Ill. 1996) (quoting *United States v. Renville*, 779 F.2d 430, 437 (8th Cir. 1985) (“The exact nature and extent of the psychological problems which ensue from child abuse often depend on the identity of the abuser.”)).

31. See, e.g., *State v. Robinson*, 735 P.2d 801, 810 (Ariz. 1987) (“[E]ffective treatment may require that the victim avoid contact with the abuser, not just to prevent future abuse, but also to facilitate recovery from past abuse.”).

32. See, e.g., *State v. Jones*, 625 So. 2d 821, 825-26 (Fla. 1993) (rejecting identification of perpetrator under medical diagnosis exception). Cf. CHRISTOPHER B. MUELLER & LAIRD C. KIRKPATRICK,

In states that restrict application of the medical diagnosis exception, prosecutors often seek to admit the statements under other exceptions, most commonly the residual exception, or, in many states, under a special exception for children's allegations of abuse. These other exceptions always require a close look at reliability. Under the residual exception, statements must have "equivalent circumstantial guarantees of trustworthiness" (that is, equivalent to other exceptions).³⁴ Under most versions of the special child abuse exception, a court must find that the statement evinces "indicia of reliability."³⁵ If the child declarant does not testify, statements admitted under either exception must be shown to enjoy "particularized guarantees of trustworthiness" to pass muster under the Confrontation Clause, because neither exception is "firmly rooted."³⁶ When an analysis of reliability is added to the usual consideration—whether a statement was made for the purpose of diagnosis or treatment—the potential suggestibility of the child comes to the fore.

At the same time that courts are taking a closer look at the reliability of young children's hearsay, researchers have focused on the suggestibility of preschoolers. Since Bruce Larson was originally convicted, a massive amount of research has been conducted on children's suggestibility. At first glance, the case appears to be a good candidate for a suggestibility defense. The child was three years old when she first accused her father of abuse.³⁷ Suggestibility researchers have shown that preschool children are particularly vulnerable to suggestive influences.³⁸ At the time the allegations surfaced, Larson and his

EVIDENCE: PRACTICE UNDER THE RULES, § 8.42 at 942 (2d ed. 1999) (identification aids in "providing social remedies" rather than medical treatment).

33. See Mosteller, *Maturation*, *supra* note 28, at 62 ("The typical reasonable person/declarant would not likely consider a statement to a doctor identifying the perpetrator of a sexual assault to be critical to proper medical treatment."); Mosteller, *Child Sexual Abuse*, *supra* note 14, at 277 n.83 ("Before identification statements . . . are included as ancillary to other statements motivated by the declarant's selfish interest, the prosecution should make a real showing that the declarant had such a subjective recognition of the overall significance of her statements.").

34. FED. R. EVID. 807.

35. See, e.g., WASH. REV. CODE ANN. § 9A.44.120 (2000) (requiring that "the time, content, and circumstances of the statement provide sufficient indicia of reliability"); MYERS, *supra* note 25, § 7.53, at 355 ("Most child hearsay exceptions are modeled on the Washington formula.").

36. See MYERS, *supra* note 25, § 7.61, at 400 ("Child hearsay exceptions are specialized residual exceptions, and are not firmly rooted."); *Swan v. Peterson*, 6 F.3d 1373, 1379 (9th Cir. 1993) ("[e]nacted in 1982, [Washington's child sexual abuse hearsay] exception is relatively new and not firmly rooted."); *State v. Robinson*, 722 P.2d 1379, 1385 (Wash. App. 1986) (stating that Washington's child sexual abuse hearsay exception "is not a firmly rooted hearsay exception"). But see Stephen J. Ceci & Richard D. Friedman, *The Suggestibility of Children: Scientific Research and Legal Implications*, 86 CORNELL L. REV. 33, 94 n.216 (2000) ("It is an unresolved question whether 'tender years' exceptions such as Washington's will be deemed to be 'firmly rooted' and so to satisfy the Confrontation Clause without a need for individualized inquiry into trustworthiness.").

37. See *State v. Larson*, 472 N.W. 2d 120, 122 (Minn. 1991).

38. See, e.g., Peter A. Ornstein & Betty N. Gordon, *The Psychologist as Expert Witness: A Comment*, in EXPERT WITNESSES IN CHILD ABUSE CASES: WHAT CAN AND SHOULD BE SAID IN COURT 237, 244 (Stephen J. Ceci & H. Hembrooke eds., 1998) ("A growing body of work indicates that young children (particularly pre-schoolers) are especially vulnerable to external influences in the form of misleading or suggestive information."); Karen J. Saywitz & Thomas D. Lyon, *Coming to Grips with Children's Suggestibility*, in MEMORY AND SUGGESTIBILITY IN THE FORENSIC INTERVIEW (M.L. Eisen, G.S. Goodman & J.A. Quas eds., forthcoming 2001) (copy on file with author). Cf. Debra A.

wife were divorced and she had custody of the child.³⁹ Suggestibility researchers emphasize the dangers of pressure from adults in abuse allegations arising after divorce.⁴⁰ The child was interviewed by a number of professionals over the course of several months.⁴¹ Suggestibility researchers warn against multiple suggestive interviews over an extended period.⁴²

On the other hand, much of the research on children's suggestibility does not appear to fit the facts of this case. In *Larson*, the mother did not initiate the examination because of suspicions of abuse, there was no evidence of coaching, and the child was more willing to talk about the abuse with the physician's assistant than with the mother. Although the child was young and thus vulnerable to leading questions, the physician's assistant was the first to question the child, there was no use of anatomical dolls,⁴³ and no peer pressure. No one argued that the physician's assistant had selectively reinforced the child's responses;⁴⁴ she asked about sexual touching but also inquired about innocent touching. There was no evidence of stereotype induction—telling the child that her father was a bad man.⁴⁵ Indeed, the child clearly showed positive feelings for her father.⁴⁶ Subsequent interviews were subject to many of these problems,⁴⁷ but the child's initial statements to the physician's assistant were clearly the most damning.⁴⁸ The defense needed to undermine the reliability of the child's first statements to argue that subsequent interviews created the allegations, rather than merely sustained them.

Poole & Michael E. Lamb, *INVESTIGATIVE INTERVIEWS OF CHILDREN: A GUIDE FOR HELPING PROFESSIONALS* 58 (1998) ("Although authors frequently comment that preschool children make more errors than older children or adults, suggestibility to misleading questions declines gradually well into the school years.").

39. See *Larson*, 472 N.W. 2d at 122.

40. See, e.g., Ceci & Friedman, *supra* note 36, at 59 ("Many experts can cite cases—in acrimonious disputes, for example—in which one parent is accused of coaching the child against the other parent.").

41. See *Larson*, 472 N.W. 2d at 122-23.

42. See, e.g., Maggie Bruck et al., *Reliability and Credibility of Young Children's Reports: From Research to Policy and Practice*, 53 AM. PSYCHOL. 136, 141 (1998) ("[C]hild witnesses are often interviewed over a prolonged period of time, and they are reinterviewed on many occasions about the same set of suspected events.").

43. See 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); Maggie Bruck, *Children's Use of Anatomically Detailed Dolls to Report Genital Touching in a Medical Examination: Developmental and Gender Comparisons*, 6 J. EXPERIMENTAL PSYCHOL.: APPLIED 74 (2000).

44. See Sena Garven et al., *More Than Suggestion: The Effect of Interviewing Techniques From the McMartin Preschool Case*, 83 J. APPLIED PSYCHOL. 347 (1998) [hereinafter Garven et al., *More Than Suggestion*]; Sena Garven et al., *Allegations of Wrongdoing: The Effects of Reinforcement on Children's Mundane and Fantastic Claims*, 85 J. APPLIED PSYCHOL. 38 (2000) [hereinafter Garven et al., *Allegations of Wrongdoing*].

45. See Michelle D. Leichtman & Stephen J. Ceci, *The Effects of Stereotypes and Suggestions on Preschoolers' Reports*, 31 DEVELOPMENTAL PSYCHOL. 568 (1995).

46. See *Larson*, 472 N.W. 2d at 122-23.

47. See *Larson*, 472 N.W. 2d at 123 (stating that therapist interviewed child with anatomically detailed dolls); see *id.* at 127 (stating that child protection specialist interviewed child with anatomically detailed dolls); *id.* (stating that police officer, who questioned child with social worker, asked "leading questions that began with the premise that 'Daddy hurt' B.").

48. See *id.* at 126 ("Of B's statements, those made to the physician's assistant clearly were key.").

What appears to be most problematic about the physician's assistant's questioning was that she repeated her question after the child's initial denial that inappropriate touching had occurred. Suggestibility researchers contend that young children will change their answer to questions if the questions are repeated within an interview.⁴⁹ A central issue in *Larson* is whether repeating questions may have led to a false allegation of abuse.

III

THE IMPORTANCE OF QUESTION REPETITION AND COMMON WISDOM REGARDING ITS DANGERS

The issue of question repetition is of special interest for several reasons. First, there are many factors that cause interviewers to repeat questions when interviewing children. In addition to repeating questions in response to apparent reluctance on the part of the child to report abuse, interviewers repeat questions to elicit additional information, to clarify ambiguous responses, and because of their uncertainty (or forgetfulness) over whether a topic was covered. Although there is some debate regarding whether the highly suggestive techniques used in several infamous day care cases and mimicked in several well-known studies are typical of mundane sexual abuse investigations,⁵⁰ question repetition is quite common. Amye Warren and her colleagues examined the transcripts of forty-two videotaped investigative interviews and found that, in ninety-five percent of the interviews, the interviewer repeated a question that had been unambiguously answered by the child in the immediately preceding portion of the interview; seven percent of all questions were repeated.⁵¹

Second, if simply repeating a question undermines the reliability of a child's testimony, this may justify expert testimony on the suggestibility of children. When suggestibility experts testify in court that children are more suggestible than adults, or that highly coercive interviewing techniques increase error, they are repeating what most lay people already believe, making it questionable whether the expert testimony is helpful to the jury.⁵² On the other hand, the ef-

49. See Stephen J. Ceci & Maggie Bruck, *Children's Testimony: Applied and Basic Issues*, in 4 HANDBOOK OF CHILD PSYCHOL. 713, 740 (I.E. Sigel & K.A. Renninger eds., 5th ed. 1998) [hereinafter Ceci & Bruck, *Children's Testimony*]; Ceci & Friedman, *supra* note 36, at 53.

50. See Ceci & Friedman, *supra* note 36, at 57-70; Lyon, *supra* note 25, at 1033-42; Amye Warren et al., *Setting the Record Straight: How Problematic are "Typical" Child Sexual Abuse Interviews?*, Paper presented at the biennial meeting of the American Psychology-Law Society (March 2000) (on file with author) (concluding that "the assumptions made by Lyon (1999) about the rarity of some of the most egregious interviewing practices (e.g. referring to what other people have said) in 'typical' interviews may be well founded," but adding several "caveats" regarding the representativeness of the interviews, the coding system, and the fact that even infrequent use of some suggestive methods is problematic); Garven et al., *More Than Suggestion*, *supra* note 44, at 354-55 ("[T]he interviewing techniques from the McMartin case should not be viewed as typical of the practice in most child protection and law enforcement agencies.").

51. Warren et al., *supra* note 50; see also Ceci & Friedman, *supra* note 36, at 66 n.168 (discussing Warren's results).

52. See Ceci & Friedman, *supra* note 36, at 100 ("One can easily accept the proposition—which Lyon supports with survey evidence—that many, even most, potential jurors understand that children

fects of mere repetition may strike jurors as surprising.⁵³ Expert input may also help courts consider the reliability of children's out-of-court statements when assessing their admissibility. The courts have some understanding of the dangers of repetition; they consider whether the child's statements are spontaneous or are the product of repeated questions in applying the excited utterance hearsay exception⁵⁴ and the residual or catch-all exception.⁵⁵ They might be surprised, however, by the magnitude of the effects of repeating questions. The case law reveals that expert opinions discussing the negative effects of repeated questions have reached the courts through expert testimony,⁵⁶ through amicus briefs, and through researchers' published writings.⁵⁷ Indeed, in *Larson*, an expert testifying before the recent research on suggestibility criticized the use of repeated questions within interviews.⁵⁸

Researchers have characterized repeating questions as a "highly suggestive" technique⁵⁹ and as a "very powerful suggestive manipulation."⁶⁰ Stephen Ceci

are more suggestible than adults, and yet recognize the value of expert evidence"); Thomas D. Lyon, *The New Wave in Children's Suggestibility Research: A Critique*, 84 CORNELL L. REV. 1004 (1999) [hereinafter Lyon, *New Wave*] (responding to Ceci and Bruck's claim that nonexperts are less likely to recognize children's weaknesses than their strengths by citing "[v]arious surveys of jury-eligible citizens" that "demonstrate that jurors are well aware of the potential for suggestibility"); 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 155, 159 (Stephen J. Ceci et al. eds., 1987) (noting that a survey of parents and college students found that the majority of respondents [77%] saw five- to nine-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 37, 38 (Stephen J. Ceci et al. eds., 1989) (providing a survey of college students showing they "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 witness of adult age (either young or old)"); A. Daniel Yarmey & Hazel P. Tressillian 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) (finding that 47% of citizens eligible to be jurors said an eight-year-old would answer questions by police or in court "the way he/she thinks the questioner wants him to," 23% said the child would say "I don't know" or refuse to respond to questions, and 30% said the child would "reply accurately").

53. See Ceci & Friedman, *supra* note 36, at 101 n.286 (arguing that many jurors will not recognize the effects of repeated questions); Steven I. Friedland, *On Common Sense and the Evaluation of Witness Credibility*, 40 CASE W. RES. L. REV. 165, 181 (1989) ("[J]urors probably fail to take [effects of repeated questions] into account.").

54. See, e.g., *Commonwealth v. Vining*, 744 A.2d 310, 317-18 (Pa. Super. Ct. 1999).

55. See, e.g., *State v. Robinson*, 735 P.2d 801, 814 (Ariz. 1987).

56. See, e.g., *United States v. Rouse*, 111 F.3d 561, 571 (8th Cir. 1997); *State v. Slane*, 1999 WL 961453, at *11 (Ohio App. 1999).

57. See, e.g., *State v. Michaels*, 642 A.2d 1372, 1377 (N.J. 1994) ("When a child is asked a question and gives an answer, and the question is immediately asked again, the child's normal reaction is to assume that the first answer was wrong or displeasing to the adult questioner") (citing Debra A. Poole & Lawrence T. White, *Effects of Question Repetition on the Eyewitness Testimony of Children and Adults*, 27 DEVELOPMENTAL PSYCHOL. 975 (1991) [hereinafter Poole & White, *Question Repetition*]); *Rouse*, 111 F.3d at 576 (Bright, J., dissenting) (citing Stephen J. Ceci & Maggie Bruck, *The Suggestibility of the Child Witness: A Historical Review and Synthesis*, 113 Psychol. Bull. 403 (1993) [hereinafter Ceci & Bruck, *Suggestibility*]) ("Repeated questions can produce a change of answers as the child may interpret the question as 'I must not have given the correct response the first time.'").

58. Appellant's Brief on Remand From the United States Supreme Court at 30-31 n.2, *State v. Larson*, 453 N.W.2d 42 (Minn. 1990) (No. C2-88-2379).

59. Ceci & Friedman, *supra* note 36, at 53.

and Maggie Bruck, two of the most prominent suggestibility researchers, report that “a number of studies, from different domains, demonstrate that when young children are asked the same question more than once within an interview, they change their answer.”⁶¹ Stephen Ceci and Richard Friedman, reviewing the suggestibility research for a legal audience, argue that preschool children are disproportionately likely to change their answers to questions such as “Did he touch you there?” merely upon repetition of the question.⁶²

Researchers have offered two reasons why children are likely to change their answers when questions are repeated. First, children believe that their first answer must have been incorrect.⁶³ Second, children assume that the adult questioner is unhappy with the answer, and they change their answer to please the adult.⁶⁴

When one turns from research to policy, the research on repeated questioning has implications for those who interview children and those who evaluate child interviews. The interviewer who weighs the advantages and disadvantages of questioning techniques must consider whether to repeat questions in order to elicit an abuse report. Suggestibility researchers implicitly agree that the disadvantages outweigh any advantages, because interviewers are warned to avoid repeating closed-ended questions within an interview.⁶⁵ A court evaluating the reliability of a child’s statement must consider the effects of repeated questions on the accuracy of the child’s report. Suggestibility researchers believe that repetition undermines the reliability of children’s statements. In discussing when a child’s hearsay statement *might* be sufficiently reliable, Ceci and Friedman posit a child who, “*without any prompting*, articulates a detailed and plausible account of abuse soon after the alleged event, and, still without prompting, consistently adheres to that account.”⁶⁶

60. Johann Endres, Claudia Poggenpohl & Christina Erben, *Repetitions, Warnings, and Video: Cognitive and Motivational Components in Preschool Children’s Suggestibility*, 4 LEGAL & CRIMINOLOGICAL PSYCHOL. 129, 133 (1999).

61. Ceci & Bruck, *Children’s Testimony*, *supra* note 49, at 740.

62. Ceci & Friedman, *supra* note 36, at 101 n.286.

63. See Robyn Fivush & April Schwarzmüller, *Say it Once Again: Effects of Repeated Questions on Children’s Event Recall*, 8 J. TRAUMATIC STRESS 555, 573 (1995) (“When asked the same question within the same interview, young children seem to interpret the repetition as an indication that their initial response was wrong, and therefore they change their answer”); Ceci & Bruck, *Children’s Testimony*, *supra* note 49, at 740 (“The first answer I gave must be wrong, that is why they are asking me the question again”).

64. See STEPHEN J. CECI & MAGGIE BRUCK, *JEOPARDY IN THE COURTROOM: A SCIENTIFIC ANALYSIS OF CHILDREN’S TESTIMONY* 120 (1995) [hereinafter CECI & BRUCK, *JEOPARDY*] (“At other times, children may change their answer to please the adult who is questioning them.”); Friedland, *supra* note 53, at 181; Philip J. Kinsler, *Children in Court: What Do We Now Know?* 24 VT. BAR J. & L. DIG. 47, 48 (Dec. 1998).

65. See Ceci & Friedman, *supra* note 36, at 86; Lynne Celandier De Sarbo, *The Danger of Value-Laden Investigation in Sexual Abuse Cases: Are Defendants’ Constitutional Rights Violated When Mental Health Professionals Offer Testimony Based on Children’s Hearsay Statements and Behaviors?*, 2 U. PA. J. CONST. L. 276, 285 (1999), citing Lucy S. McGough & Amye R. Warren, *The All-Important Investigative Interview*, 45 JUV. & FAM. CT. J. 13 (1994).

66. Ceci & Friedman, *supra* note 36, at 96 (emphasis added).

IV

A CLOSER LOOK: THE EFFECTS OF REPETITION

In applying research on repeated questions to a particular case, one must move beyond general statements about how repetition undermines children's reliability. Application of research to real cases requires a careful assessment of the research methodology, the participants' characteristics, and the research results, as well as an evaluation of the interviewing technique and the child interviewed. Experts must be forthright about the "boundary conditions" of the research, which potentially limit the extent to which the research findings can be applied to a particular case.⁶⁷ Social scientists commonly refer to these questions as issues of external validity, which concerns the extent to which research results apply to other people, in other situations, and at other times.

A court assesses the boundary conditions or external validity of research when it decides whether an expert's proffered testimony sufficiently "fits" the facts of a case to be helpful to the jury, or whether it is helpful to the judge in assessing the admissibility of evidence. Exact fit is not required, and experts may exercise their best judgment as to how the research applies. Research that is inapplicable to a particular case, however, is likely to mislead the finder of fact and should be excluded.⁶⁸

A number of factors influence the potential effects of question repetition upon children's reliability. Here, the focus will be on the most important factors, which include whether questions were repeated within an interview (or over repeated interviews), the types of questions asked, the child's age, and the child's memory of the event. The leading research reveals that question repetition does not inevitably increase error.

A. Repeating Questions Versus Repeating Interviews

In their overview of the literature, Ceci and Friedman discuss five studies "focusing on repeated questions" that find high rates of error in response to abuse-related questions.⁶⁹ A careful reading of Ceci and Friedman's paper reveals that only one of those studies analyzed the effects of repeating questions *within* an interview, rather than *across* interviews.⁷⁰ The distinction between re-

67. See CECI & BRUCK, *JEOPARDY*, *supra* note 64, at 273 ("In short, we urge expert witnesses to review the full corpus of relevant scientific work, describing the magnitude of errors, the inconsistencies within and across studies, and the boundary conditions that might limit any generalization from the science to the case at bar); Ornstein & Gordon, *supra* note 38, at 242 ("boundary conditions' may interfere with general statements about the sequelae of abuse or the suggestibility of children, requiring the expert to say that 'X may hold under certain conditions, but Y may be true in other situations.'").

68. See generally Thomas D. Lyon, *Expert Testimony on the Suggestibility of Children: Does it Fit?*, in *CHILDREN AND THE LAW: SOCIAL SCIENCE AND POLICY* (B.L. Bottoms et al. eds.) (forthcoming 2002, available at <http://hal-law.usc.edu/users/tlyon/articles/bottoms.pdf>) [hereinafter Lyon, *Expert Testimony*].

69. Ceci & Friedman, *supra* note 36, at 55.

70. When the authors specifically refer to "repetition of questions within the same interview," they cite one of the five studies: the 1991 study by Poole and White, discussed at length *infra* text accompa-

peating questions within interviews rather than across interviews is important.⁷¹ Repeated questions *within* interviews may lead to error because of the child's perception that a different response is expected. Repeating questions *across* interviews may also lead to increased error, but for different reasons: Children forget over time, so that later interviews are more error-filled, and children may confuse what actually occurred with their responses in earlier interviews. On the other hand, repeating questions across interviews may decrease error. Repetition is a form of rehearsal, which strengthens memory, and children may recall new details during subsequent interviews (something memory researchers call "reminiscence").⁷² Therefore, errors across interviews do not tell us whether children will err within an interview.

B. The Leading Research: Poole and White

The study that speaks directly to the issue of repeated questions within interviews was conducted by Debra Poole and Lawrence Wright.⁷³ As Ceci and Friedman summarize the study, "Poole & White interviewed some four-, six-, and eight-year-olds, and adults immediately after a staged encounter with a man, and again one week later."⁷⁴ In each interview, questions were repeated twice, so that they were asked a total of three times over one interview, and six times over two interviews. Summarizing the findings, Ceci and Friedman note that repetition of the question "Did the man hurt Melanie?" increased the likelihood of a "false affirmative answer" among the four-year-olds. Indeed, sixty percent of the subjects to whom the question was asked six times over two interviews ultimately answered "yes."⁷⁵

It is immediately apparent that the Poole and White study differs from the situation in *Larson*. Poole and White examined repetition over two interviews, rather than one, and asked the question six times, rather than twice. Nevertheless, a sixty percent rate of false affirmation is alarming, particularly since the question is arguably "abuse-related."⁷⁶ Moreover, the study is cited more generally as demonstrating age differences in the consistency of responses to repeated questioning, with young children most inclined to change their answers.⁷⁷

Should this study lead the court to find the child's statements unreliable in *Larson*? Or, if the court admits the statements, does the study justify expert

nying notes 73-84. They also cite the two-year followup to the Poole and White study, discussed *infra* text accompanying notes 85-91. *Id.* at 53.

71. See, e.g., Fivush & Schwarzmüller, *supra* note 63, at 568 ("questions asked repeatedly during the same interview session may be perceived differently than the same question asked again on different recall occasions").

72. See *id.* at 556 (describing how repeated interviews might improve memory).

73. Poole & White, *Question Repetition*, *supra* note 57. This is perhaps the most oft-cited study regarding question repetition within interviews.

74. Ceci & Friedman, *supra* note 36, at 55.

75. *Id.* at 52.

76. *Id.* at 57.

77. See Ceci & Bruck, *Children's Testimony*, *supra* note 49, at 740; Poole & Lamb, *supra* note 38, at 56; Fivush & Schwarzmüller, *supra* note 63, at 570.

testimony that repeated questioning will lead to false affirmations about harmful touching? Several important details about the Poole and White study suggest negative answers to both questions.

First, a “yes” response to “Did the man hurt Melanie?” is not a “false affirmation.” Rather, it is entirely reasonable to conclude that the man did hurt Melanie. Poole and White intentionally staged the event so that “whether Melanie had actually been injured or was simply discouraging” the man and “whether the interaction was predominately antagonistic” were “ambiguous features.”⁷⁸ Second, children were no more likely than adults to view the interaction as harmful. The researchers found no evidence that “younger subjects were more or less prone to report that injury had occurred or that inappropriate behavior was witnessed.”⁷⁹ In part, this finding emphasizes the ambiguity of the interaction, because adults were just as unsure as children about what had occurred. Third, subjects tended to minimize the extent to which a harmful interaction occurred. The researchers believed that the question, “Did the man ask nicely for the pen?” was unambiguous “because the assistant inappropriately grabbed the pen from the subject.” Nevertheless, “many subjects were reluctant to say anything unflattering about the stranger.”⁸⁰ Poole and White suggest that either subjects were making judgments based on the overall actions of the man, or that “there is implicit social pressure involved in reporting antisocial behavior.”⁸¹ Subjects’ denial that the man acted anti-socially, either by grabbing the pen or hurting the woman, is evidence of false denial rather than false affirmation. Hence, the study raises concerns that subjects might conceal wrongdoing by another person because of their reluctance to accuse that person of doing wrong.

The researchers found that younger children (four-year-olds) were more likely to change their responses to yes/no questions than were older children and adults.⁸² Four-year-olds changed their responses eight to sixteen percent of the time when questions were repeated, whereas six-year-olds and older subjects did so only one to two of the time.⁸³ Given the ambiguity of the interac-

78. Poole & White, *Question Repetition*, *supra* note 57, at 977.

79. *Id.* at 981. Specifically, adult subjects who answered the question six times were no less likely than children who answered the question six times to claim that the man had hurt Melanie. A careful reader would detect this fact in Ceci and Friedman’s paper. In the text, the authors refer to the tendency of the “four-year-olds” to give a false affirmative answer, Ceci & Friedman, *supra* note 36, at 56, whereas in a footnote they report percentages and note that they apply “[r]egardless of the subjects’ age and gender.” *Id.* at 56 n.115.

80. Poole & White, *Question Repetition*, *supra* note 57, at 984.

81. *Id.* at 984.

82. Yes/no questions are questions that can be answered “yes” or “no,” for example, “Did the man ask nicely for the pen?”

83. Poole & White, *Question Repetition*, *supra* note 57, at 981. The authors report that “[a]cross both repeat and single-delayed conditions, 24% of the four-year-olds were inconsistent on at least one yes/no question, compared with only 6% of the older subjects.” *Id.* at 980-81. This figure is necessarily higher than the likelihood of inconsistency per question, because children were asked two yes/no questions. Moreover, this figure appears to include not only inconsistencies due to repetition but also inconsistencies across the two interviews.

tion, however, the researchers cautiously concluded that “[m]ore research is needed to determine whether the repetition of yes-no questions is generally problematic for preschoolers or whether inconsistencies will be apparent only for minor details or ambiguous interactions.”⁸⁴

Poole and White returned two years later and re-interviewed many of the original subjects.⁸⁵ As before, questions were repeated; this time subjects were asked each question a total of three times. Not surprisingly, many children remembered little of the one-minute interaction they had witnessed two years earlier.⁸⁶ Summaries of the follow-up study have emphasized the age differences in errors: Younger children performed less well than older children and adults,⁸⁷ even in response to open-ended questions; one-fifth of the children confused which actors performed certain actions,⁸⁸ and children were more likely to speculate in response to some questions than adults. Yet, for our purposes, the most important question is whether younger children were more likely than older children and adults to change their answers in response to repeated questions. The answer is no. Across both studies, subjects changed their answer when the question was repeated about five percent of the time.⁸⁹ The researchers report being “extremely surprised that question repetition was not associated with decreased accuracy, particularly given the fact that the youngest children were only four years old when they witnessed the event.”⁹⁰

The follow-up study found further evidence that subjects, particularly children, are reluctant to label behavior as anti-social. Although they were no more likely than adults to change their responses within the interview, children were more likely than adults to change their responses across interviews, that is, from the interviews shortly after the original event to the interviews two years after the event. The direction of their change was significant: about ninety percent of the time, children who changed their responses “initially reported that the man behaved badly but later claimed that he had not.”⁹¹ The researchers found evidence that these children had forgotten the interaction, which suggests that a child who does not remember an event is not prone to report that event as anti-social, even if questions are repeated.

The two Poole and White studies provide little evidence that non-abused children are likely to fabricate abuse allegations in response to repeated yes/no

84. *Id.* at 984.

85. Debra A. Poole & Lawrence T. White, *Two Years Later: Effects of Question Repetition and Retention Interval on the Eyewitness Testimony of Children and Adults*, 29 DEVELOPMENTAL PSYCHOL. 844 (1993) [hereinafter Poole & White, *Two Years Later*].

86. See Poole & White, *Question Repetition*, *supra* note 57, at 977 (“the interaction between assistants was very brief, lasting less than a minute”).

87. See Ceci & Bruck, *Suggestibility*, *supra* note 57, at 420.

88. See CECI & BRUCK, JEOPARDY, *supra* note 64, at 108-09.

89. See Poole & White, *Two Years Later*, *supra* note 83, at 847 (“the mean consistency score within a session was 1.91 out of a maximum score of 2”).

90. *Id.* at 851. See also *id.* at 849 (“There was no evidence that testimony became proportionately less accurate with repetition in Study 2.”).

91. *Id.* at 847-48.

questions. Although the youngest children in the first study were more likely than older subjects to change their answers when questions were repeated, the follow-up study found no distinction among age groups. Even in the first study, the youngest subjects were no more likely than adults to claim after repeated questions that something anti-social had occurred. Rather, the studies suggest that children, like adults, are reluctant to accuse others of wrongdoing, at least when the interaction is ambiguous or difficult to recall. These findings do not argue against the reliability of an abuse allegation in response to a repeated question.

C. The Classic Research: Piaget and His Detractors

Poole and White's work is not the only research upon which reviewers have relied in arguing that repetition breeds inconsistency and error. Indeed, the notion that repeated questions elicit changes in children's responses is a long-standing one, dating back to research in the 1970s criticizing some of Jean Piaget's classic studies. Researchers argued that some of Piaget's findings were attributable to the fact that he repeated questions, eliciting inconsistent responses from young children. Suggestibility researchers have widely cited the research challenging Piaget's results to support the claim that repetition increases errors.⁹² Careful scrutiny of that research, however, reveals that there are important limits to the effects of repetition.

In order to understand the research, it is necessary to understand a little about Piaget's developmental theories. Piaget argued that young children fail to conserve number, mass, and volume. For example, with respect to the conservation of number, Piaget argued that young children do not understand that if one changes the spacing of a row of objects, the number of objects does not change.⁹³ In a number conservation study, an experimenter shows a child two rows of objects that are equal in number and spaced similarly. The experimenter asks the child if one row has more objects or if they are both the same.⁹⁴ The experimenter then spreads out one of the rows and repeats the question. Younger children tend to change their answers, first responding that the rows are equal, but the—after the experimenter's manipulation—responding that the transformed row has more objects.

92. See Ceci & Bruck, *Suggestibility*, *supra* note 57, at 419; Amina Memon & Rita Vartoukian, *The Effects of Repeated Questioning on Young Children's Eyewitness Testimony*, 87 BRIT. J. PSYCHOL. 403 (1996); Poole & White, *Question Repetition*, *supra* note 57, at 975; Debra A. 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, 36 (Maria S. Zaragoza et al. eds., 1995) [hereinafter Poole & White, *Tell Me Again*]; Christine M. Ricci & Carole R. Beal, *Effect of Questioning Techniques and Interview Setting on Young Children's Eyewitness Memory*, 6 EXPERT EVIDENCE 127, 128 (1998).

93. See JEAN PIAGET, *THE CHILD'S CONCEPTION OF NUMBER* (Caleb Gattegno & F.M. Hodgson trans., Hermitage Press 1952).

94. The question could be phrased as a yes/no question ("Do these rows have the same number?") or as a forced-choice question ("Are there the same number in both rows or does one have more?").

Piaget's interpretation of his results was challenged by researchers who argued that a child in the conservation experiment might "feel that the second question itself seems to suggest that a new judgment is in order, with the result that he changes his answer."⁹⁵ Susan Rose and Marion Blank found that if first graders were asked whether the rows were equal only once—after the transformation occurred, the number of conservation errors was cut in half.⁹⁶ Other researchers replicated Rose and Blank's findings on number conservation for the other conservation tasks—volume, mass, and length.⁹⁷ Moreover, Michael Siegal and his colleagues showed that children interpreted the repetition of the question as a request for a different answer. Siegal's research team found that when four- to six-year-old participants watched puppets making errors on the standard conservation task, they tended to attribute those errors to the puppet's desire "to please the grownup" rather than what "the puppet really thought was true."⁹⁸

At first glance, the implications of the research for children's eyewitness memory seem clear: Merely repeating the conservation question appears to elicit errors in young children. If children err when responding about objects they are currently viewing, then surely they will err in responding about events they are trying to remember.

Consider, however, the context in which repetition occurs in the conservation experiment. The researcher asks the child a question, transforms the materials, and then asks the question again. The repetition is coupled with transformation of the objects. If repeating the question tells a child that her first answer was incorrect, repetition alone (either before or after transformation) should have equally deleterious effects. Several researchers have found that this is not the case. Studying three- to four-year-olds, Graeme Halford and Frances Boyle asked two number-conservation questions without manipulating the objects between questions; the experimenter asked the question, dropped something on the floor, and then repeated the question.⁹⁹ The authors found that "the preschool children seem to have no difficulty maintaining their judgment when they are requested to repeat it without seeing a transformation."¹⁰⁰ Whereas children changed their answers about half the time in the standard experiment, they did so 12.5% of the time when the question was simply re-

95. Susan A. Rose & Marion Blank, *The Potency of Context in Children's Cognition: An Illustration Through Conservation*, 45 CHILD DEV. 499, 500 (1974).

96. *Id.* ("[E]rrors on conservation were more than halved on the one-judgment task relative to the other two tasks.")

97. See Judith Samuel & Peter Bryant, *Asking Only One Question in the Conservation Experiment*, 25 J. CHILD PSYCHOL. & PSYCHIATRY & ALLIED DISCIPLINES 315 (1984) (conservation of mass, number, and volume); Michael Siegal, Lorraine J. Waters & Leigh S. Dinwiddy, *Misleading Children: Causal Attributions For Inconsistency Under Repeated Questioning*, 45 J. EXPERIMENTAL CHILD PSYCHOL. 438 (1988) (conservation of number and length).

98. Siegal, Waters & Dinwiddy, *supra* note 97, at 445.

99. Graeme S. Halford & Frances M. Boyle, *Do Young Children Understand Conservation of Number?*, 56 CHILD DEV. 165, 170 (1985). The experimenters asked forced-choice conservation questions. *Id.*

100. *Id.* at 171.

peated.¹⁰¹ Studying six-year-olds, Neilson and colleagues asked two number conservation questions either before manipulating the objects or afterward.¹⁰² While children changed their answers about half the time in the standard experiment, six percent did so when the question was repeated without the transformation.¹⁰³ Repetition alone does not clearly convey to children that their answer is wrong; rather, something else must occur to suggest that a different answer is appropriate.¹⁰⁴

The history of Piaget's original research and the research it spawned teaches important lessons about the application of research to real-world cases. Clearly, young children can be led by the interviewing context to change their answers, but repetition does not appear to be inherently leading. Rather, one must examine the context in which repetition occurs. In one review of this and more recent research, Poole and White note that "[i]t is safe to assume that repetition effects will not be uniform but will be dependent on variables that alter witnesses' interpretations of repeated requests."¹⁰⁵ For legal decision-makers analyzing children's reports of abuse, the nature of those variables is as important as more general observations regarding the risks of repetition.

D. The Type of Questions and the Age of the Child

If one examines the entire corpus of research on repeated questions, the findings may appear inconsistent. Some studies find age differences only for younger children, whereas other studies find age differences across childhood. It is possible to derive a general rule, however. As questions become less leading, repetition has less of an effect, and even young children's responses are quite consistent; as questions become more leading, repetition has larger effects, and differences between children emerge at older ages. In assessing the effects of repetition, it is important to consider both the types of questions asked and the age of the child.

Preschool children are particularly suggestible.¹⁰⁶ In *Larson*, the child was not quite four years old when she was questioned by the physician's assistant, making research on preschool children relevant to that case. Ceci and Friedman warn, however, that the results of suggestibility research with preschool

101. *Id.*

102. See Irene Neilson, Julie Dockrell & Jim McKechnie, *Does Repetition of the Question Influence Children's Performance in Conservation Tasks?*, 1 BRIT. J. DEVELOPMENTAL PSYCHOL. 163 (1983).

103. *Id.* at 166-67. The authors asked forced-choice questions.

104. Not only is it necessary to manipulate the objects between questions, but it is necessary that the objects be manipulated by the experimenter. McGarrigle and Donaldson, studying eighty four- to six-year-olds, showed that if someone other than the experimenter manipulated the objects (a "naughty" teddy bear), children were not influenced by repetition of a forced-choice question ("Are there more here or more here or are they both the same number?"). See James McGarrigle & Margaret Donaldson, *Conservation Accidents*, 3 COGNITION 341 (1974).

105. Poole & White, *Tell Me Again*, *supra* note 92, at 37. See also Poole & White, *Question Repetition*, *supra* note 57, at 976 ("[T]here is no uniform effect of question repetition on children's performance.").

106. See discussion *supra* note 38 and accompanying text.

children is of questionable applicability to older children.¹⁰⁷ Older children are much more resistant to suggestive questions, making it necessary to increase the suggestiveness of interviewing to elicit errors and demonstrate age differences in performance.

The precise meaning of "suggestive question" is unclear.¹⁰⁸ Perhaps it is most useful to think of questions along a continuum: On one end of the continuum the child provides the details, whereas on the other end the interviewer provides the details. If the child is asked to supply all the details, the questions are less leading. As the questions move toward the interviewer-supplied end of the continuum, they become more leading. At the extreme end of the interviewer-supplied continuum would be statements by the interviewer *telling* the child what occurred.

Questions that ask for free recall are the least leading. With free recall, the interviewer asks something such as "What happened?", letting the child supply all the details. With recognition questions, the interviewer provides choices and the child selects the correct choice. Recognition questions include yes/no questions and forced-choice questions, whereby the child chooses one of the offered responses (for example, "Was it A or B?"). In either case, the interviewer supplies details that the child confirms or rejects.

Between free recall and recognition questions are cued recall questions, which focus the child's attention on a particular type of acceptable answer, but allow the child to fill in the details. These include wh- questions (questions that begin with "what", "where", "when", "who", "why", or "how"), which are often classified as "general" or "specific." As wh- questions become more specific, the interviewer supplies more of the details. For example, compare "what was the man wearing?" (more general) with "what color were the man's shoes?" (more specific). Note that in the latter example, the interviewer assumes a detail (the man was wearing shoes).

When Poole and White analyzed shifts in responses across repeated yes/no questions, they found that four-year-olds were more likely to change their answers than were older subjects. They also found, however, that school-age children were remarkably resistant to the effects of repetition; indeed, six-year-olds performed as well as adults. Similarly, upon questioning the same subjects two years after the event, Poole and White were "extremely surprised" to find no age differences.¹⁰⁹ By this time the youngest subjects were six years old, leading the researchers to conclude that, across the two studies, "[s]ubjects 6-years-old [sic] or older were able to maintain good within-session consistency on yes-no questions."¹¹⁰ Recall also that children were no more likely than adults to ulti-

107. Ceci & Friedman, *supra* note 36, at 62 ("[t]he preschool research is of substantial use only in cases involving preschool children.").

108. See Lyon, *New Wave*, *supra* note 52, at 1037-42; Ceci & Friedman, *supra* note 36, at 62-63.

109. Poole & White, *Two Years Later*, *supra* note 85, at 851.

110. *Id.*

mately conclude (after a total of six repetitions) that the ambiguous interaction between the man and woman was harmful.¹¹¹

When yes/no questions are combined with specific wh- questions (for example, “What was she carrying when she came in?” and “What color badge was he wearing?”), children will sometimes change their responses with repetition, but two studies have found no age differences among school-age children in their tendency to do so.¹¹² That is, younger children are less accurate than older children (their memories are poorer), but are no more susceptible to the effects of repeated questions.

Poole and White also asked children several general wh- questions. The open-ended questions were general wh- questions. They were “general” because they called for narrative answers. They were wh- questions because they used “who, what, where, when, why, or how” (for example, “What did he look like?”, “Tell me what happened when he came into the room,” and “Why did he come into the room?”). The researchers found no age differences in response to these questions.¹¹³ Poole and White’s two-year followup study obtained the same results.¹¹⁴ Similar results were obtained in two field studies by Michael Lamb and his colleagues, who analyzed contradictions in children’s responses in real-world interviews.¹¹⁵

How can one apply the results of this research to the child in *Larson*, or to other sexual abuse allegations? The physician’s assistant might have asked “Has anyone else touched you there?” or “Who has touched you there?” or, according to the defense, “Did your daddy touch you there?” These are yes/no or specific wh- questions. The research suggests that preschool children are more inclined to change their responses when such questions are repeated than are grade school children, but by age six children are about as resistant as older children.

Research by Michael Lamb and others examining real-world sexual abuse interviews has found that yes/no questions, forced-choice questions, and specific

111. See *supra* text accompanying note 79.

112. See Memon & Vartoukian, *supra* note 92, at 410 (finding no significant correlation between age and decreased accuracy in response to repeated closed-ended yes/no and wh- questions among five- to seven-year-olds, although repetition did not decrease accuracy across age groups); Stephen Moston, *The Suggestibility of Children in Interview Studies*, 7 FIRST LANGUAGE 67, 75 (1987) (finding no significant relation between age and decreased accuracy in response to repeated questions among six- to ten-year-olds).

113. Poole & White, *Question Repetition*, *supra* note 57, at 980 (“Inconsistencies or retractions of prior testimony were rare with open-ended questions, despite the fact that some subjects answered the same question set six times.”).

114. Poole & White, *Two Years Later*, *supra* note 85, at 849 (“There was no evidence that testimony became proportionally less accurate with repetition in Study 2.”).

115. They found that such contradictions were *never* elicited by repeated “invitations,” which are questions that elicit an “open-ended response from the child” and do “not delimit the child’s focus except in the most general way.” Michael E. Lamb & Angèle Fauchier, *The Effects of Question Type on Self-Contradictions by Children in the Course of Forensic Interviews*, APPLIED COGNITIVE PSYCHOL. (forthcoming) (manuscript at 20, on file with author); Yael Orbach & Michael E. Lamb, *The Relationship Between Within-Interview Contradictions and Eliciting Interviewer Utterances*, 25 CHILD ABUSE & NEGLECT 323, 327-28 (2001).

wh- questions predominate.¹¹⁶ Moreover, any interviewer's claim that he or she "only asked open-ended questions" is subject to doubt, given interviewers' documented difficulties in accurately recalling the form of their questions.¹¹⁷ Accordingly, it is fair to assume that most repeated questions will be quite specific questions.

Researchers have more difficulty uncovering age differences among older children in response to repeated questions, and attempt to do so by increasing the suggestiveness of the questions. Yes/no questions can be made more suggestive by turning them into "tag questions" or "negative-term" questions. Tag questions are yes/no questions in which the question is expressed as a statement followed by a phrase akin to "isn't that true?" (for example, "Michaela also let other children play with the duck, didn't she?"). Negative-term questions are yes/no questions into which a "not" is inserted (for example, "Did you see a gun?" becomes a negative-term question when phrased "Didn't you see a gun?"). Some research finds that young children are more likely to acquiesce to tag questions and negative-term questions than to yes/no questions.¹¹⁸ In the Bonn Test of Statement Suggestibility, the interviewer reads the child a story and then asks a series of yes/no, tag, and forced-choice questions (in which neither response is correct).¹¹⁹ Some of the questions are repeated immediately after they are first asked,¹²⁰ and the repeated question is prefaced with a statement such as "Are you sure?" or "Listen closely to this question again."¹²¹ Because the interviewer reads the story to the child, he or she knows the contents of the story, increasing the suggestiveness of the questions.¹²² One of three studies ex-

116. Michael E. Lamb et al., *Effects of Age and Delay on the Amount of Information Provided by Alleged Sex Abuse Victims in Investigative Interviews*, 71 CHILD DEV. 1586, 1590 (2000).

117. See Maggie Bruck, Stephen J. Ceci & Emmett Francoeur, *The Accuracy of Mothers' Memories of Conversations With Their Preschool Children*, 5 J. EXPERIMENTAL PSYCHOL.: APPLIED 89 (1999); Michael E. Lamb, Yael Orbach, Kathleen J. Sternberg et al., *Accuracy of Investigators' Verbatim Notes of Their Forensic Interviews with Alleged Child Abuse Victims*, 24 L. & HUM. BEHAV. 699 (2000); Amye R. Warren & Cara E. Woodall, *The Reliability of Hearsay Testimony: How Well Do Interviewers Recall Their Interviews with Children*, 5 PSYCHOL. PUB. POL'Y & L. 355 (1999).

118. See, e.g., Jemma Greenstock & Margaret-Ellen Pipe, *Interviewing Children About Past Events: The Influence of Peer Support and Misleading Questions*, 20 CHILD ABUSE & NEGLECT 69 (1996) (finding that five- to seven-year-olds but not eight- to ten-year-olds acquiesce more to tag questions than to yes/no questions); Bruck et al., *supra* note 42, at 139 (preschoolers more likely to acquiesce to negative-term questions than to yes/no questions).

119. See Johann Endres, *The Suggestibility of the Child Witness: The Role of Individual Differences and Their Assessment*, 1 J. CREDIBILITY ASSESSMENT & WITNESS PSYCHOL. 44, 54 (1997).

120. See *id.* at 50-51 ("If . . . an almost literal repetition of the question follows an answer immediately, this will unmistakably express the questioner's discontent with the first answer and his demand that the answer be changed."). On the other hand, repeating a question immediately reduces the likelihood that the child has forgotten her previous response.

121. Cf. Warren et al., *supra* note 50 (example of a repeated question is: "I: Did he ever touch you anywhere? C: No. I: You sure? C: (Nods head yes).").

122. See James M. Lampinen & Vicki L. Smith, *The Incredible (and Sometimes Incredulous) Child Witness: Child Eyewitnesses' Sensitivity to Source Credibility Cues*, 80 J. APPLIED PSYCHOL. 621 (1995); Michael P. Toglia et al., *The Suggestibility of Children's Memory: A Social-psychological and Cognitive Interpretation*, in DEVELOPMENT OF LONG-TERM RETENTION 217 (Mark L. Howe, Charles J. Brainerd & Valerie F. Reyna eds., 1992).

aming children from four to ten years of age has found age differences in children's tendency to change their responses when questions are repeated.¹²³

Some researchers have made repetition more leading by prefacing the repeated questions with the warning that the child has made a number of mistakes.¹²⁴ Doing so increases suggestiveness in two ways. First, telling the child she has made mistakes indicates that the interviewer is knowledgeable.¹²⁵ Second, it explicitly pressures the child to change her response. Sena Garven, James Wood, and Roy Malpass have shown that if an interviewer responds "You're not doing good" to every negative response, and "Great. You're doing excellent now" to every positive response, five- to seven-year-old children will assent to most yes/no questions, even about fantastic events, and will elaborate on those responses and repeat their claims during a subsequent interview.¹²⁶ Seven studies have examined subjects' susceptibility to repeated questions using a version of Gudjonsson's Suggestibility Scales, in which the interviewer reads the subject a story and asks twenty questions. The interviewer then warns the subject that he or she made a number of mistakes and repeats the twenty questions. Fifteen of the questions are called "suggestive": they are either yes/no questions that are correctly answered "no," or forced-choice questions (for example, "Did the woman have one or two children?") for which neither answer is correct. The remaining five questions are yes/no questions that are correctly answered "yes." Subjects' tendency to change their responses with repetition of the "suggestive" questions is called "shift," and five of the seven studies have found age differences in "shift" extending from seven years of age to adulthood.¹²⁷

123. See Ingrid Candel, Harald Merckelbach & Peter Muris, *Measuring Interrogative Suggestibility in Children: Reliability and Validity of the Bonn Test of Statement Suggestibility*, 6 PSYCHOL. CRIME & L. 61, 67-68 (2000) (finding a significant relation between age and shift in response to repeated questions among four- to ten-year-olds). Cf. Endres, *supra* note 119, at 55 (finding no significant correlation between age and shift in response to repeated questions among four- to ten-year-olds); Endres, Poggenpohl & Erben, *supra* note 60, at 135 (same; four- to seven-year-olds).

124. See, e.g., Amye Warren, Katherine Hulse-Trotter & Ernest C. Tubbs, *Inducing Resistance to Suggestibility in Children*, 15 L. & HUM. BEHAV. 273, 278 (1991) (noting that subjects "were informed that they had not performed very well and that it would be necessary to go through the questions once more.").

125. See *id.* at 284 ("One major drawback of our study in attempting to generalize to actual legal situations is the fact that our interrogator knew more about the event than the 'witnesses' themselves.").

126. Garven et al., *Allegations of Wrongdoing*, *supra* note 44.

127. See A.K. Buhrman & A.R. Warren, *Reducing Suggestibility in Normal and Learning Disabled Child Witnesses*, Paper presented at the Third International Conference on Practical Aspects of Memory, College Park, Maryland (Aug. 1994) (on file with author) (noting that seven-year-olds had higher scores on shift than fifteen-year-olds); Gudbjorg Danielsdottir, Sigrun Sigurgeirsdottir, Helma R. Einarisdottir & Erlendur Haraldsson, *Interrogative Suggestibility in Children and Its Relationship with Memory and Vocabulary*, 14 PERSONALITY AND INDIVIDUAL DIFFERENCES 499 (1993) (studying six-year-olds, eight-year-olds, ten-year-olds, and twelve-year-olds and finding no age differences in shift); G. GUDJONSSON, *THE PSYCHOLOGY OF INTERROGATIONS, CONFESSIONS & TESTIMONY* 143 (1992) (discussing two studies and finding eleven- to sixteen-year-old boys had higher scores on shift than adults); Lucy A. Henry & Gisli H. Gudjonsson, *Eyewitness Memory and Suggestibility in Children with Mental Retardation*, 104 AM. J. MENTAL RETARDATION 491, 499 (1999) (finding no difference in shift between seven-year-olds and eleven-year-olds, but mentally retarded eleven-year-olds score higher on

What about children younger than four years? The research on repeated questions has focused on children four years or older, whereas researchers have found dramatic differences between three- and four-year-olds on tasks that have implications for children's performance as witnesses.¹²⁸ Because the child in *Larson* was just turning four when she first alleged abuse, she might have been especially vulnerable to repetition. Matthew Scullin and Stephen Ceci tested the Video Suggestibility Scale for Children ("Video SSC") with three- to five-year-old children.¹²⁹ Children watched a five-minute video about a birthday party. They were subsequently asked eighteen yes/no questions. Twice they were told, "You missed a few of the questions. Let's go through them again and see if you can do better this time," and questions were repeated.¹³⁰ One unexpected finding was that the three-year-olds changed their answers in response to negative feedback *less* often than did the four- and five-year-olds.¹³¹

Review of the research on repeated questions reveals the importance of paying close attention to the types of questions asked and the ages of the children tested. When interviewers ask yes/no and wh- questions, preschool children are more inclined to change their responses than are older children, but even young grade school children are remarkably consistent. When interviewers move to more suggestive techniques, repetition has a stronger effect, and age differences appear among older children. The Scullin and Ceci study, however, reveals that the relationship between age and performance is not a simple one. It may be that children initially become more vulnerable to external pressure as they approach grade school, but that this vulnerability lessens with age.

Sometimes suggestibility effects may be due to the type of question asked rather than to repetition. Reviewers often cite the work of William Cassel and colleagues as demonstrating that "children will change their answers to repeated forced-choice questions."¹³² The researchers make it quite clear, however, that they did not simply repeat questions, but asked *different* questions. In

shift than seven-year-olds); G. Richardson, G.H. Gudjonsson & T.P. Kelly, *Interrogative Suggestibility in an Adolescent Forensic Population*, 18 J. ADOLESCENCE 211 (1995) (finding adolescents score higher on shift than adults); Warren, Hulse-Trotter & Tubbs, *supra* note 124 (studying seven-year-olds, twelve-year-olds and adults and finding all groups significantly different on shift, and finding that seven-year-olds were more likely to change responses to "non-suggestive" yes/no questions, whereas twelve-year-olds and adults performed comparably).

128. See Saywitz & Lyon, *supra* note 38.

129. Matthew H. Scullin & Stephen J. Ceci, *A Suggestibility Scale for Children*, 30 PERSONALITY & INDIVIDUAL DIFFERENCES 843 (2001).

130. *Id.* at 847-49.

131. See *id.* at 853. The authors speculated that three-year-olds might have had difficulty holding in mind what they had responded and what the interviewer wanted them to respond. See *id.*

132. Garven et al., *More Than Suggestion*, *supra* note 44, at 349 (citing William S. Cassel et al., *Developmental Patterns of Eyewitness Responses to Repeated and Increasingly Suggestive Questions*, 61 J. EXPERIMENTAL CHILD PSYCHOL. 116 (1996)); Ceci & Bruck, *Suggestibility*, *supra* note 57, at 419 ("42% of kindergarten children changed their minds on repeated questioning.") (citing W.S. Cassel & D.F. Bjorklund, *Age Difference and Suggestibility of Eyewitnesses*, in *Children's Memory for Real World Events: Implications for Testimony*, Symposium conducted at the annual Conference for Human Development (Apr. 1992), subsequently published as William S. Cassel & David F. Bjorklund, *Developmental Patterns of Eyewitness Memory and Suggestibility: An Ecologically Based Short-Term Longitudinal Study*, 19 L. & HUM. BEHAV. 507 (1995)).

one study, for example, the authors found that kindergartners were more likely than older children and adults to change their answers when asked “increasingly suggestive” tag questions.¹³³ The poor performance of the kindergarten children could be attributed to the form of the second question, rather than repetition.¹³⁴

Although there was no evidence in *Larson* that the physician’s assistant went beyond repeating a yes/no question, critics of interviewing will point out that in surveys of real-world interviews, about ten percent of the questions are “suggestive,” which includes tag-questions and questions that presuppose information not provided by the child (“suppositional” questions).¹³⁵ Moreover, Warren, Garven, Walker, and Woodall found that two-thirds of all investigative interviews they reviewed contained at least one “negative consequence,” which was defined as “[c]riticizing or disagreeing with a child’s statement, or otherwise indicating that the statement is incomplete, inadequate, or disappointing.”¹³⁶ Overall, four percent of the interviewers’ statements were categorized as “negative consequences.”¹³⁷ Hence, some of the repeated questions in real-world interviews are going to be more coercive than, for example, the yes/no questions in the Poole and White studies.

The crux of the debate is whether real-world interviews rise to the level of coerciveness employed in the studies that produce dramatic suggestibility effects.¹³⁸ Ceci and Friedman argue that a single leading question can taint an interview,¹³⁹ but do not cite any research for this proposition. Research docu-

133. Cassel et al., *supra* note 132, at 127. See also Cassel & Bjorklund, *supra* note 132 (first question was a tag question suggesting the correct answer (“positive leading”), whereas the second question was a tag question suggesting an incorrect answer (“negative leading”)).

134. Ceci & Bruck, *Children’s Testimony*, *supra* note 49, at 740 (discussing Cassel & Bjorklund: “If [the children studied] did not fall sway to the lead, then they were asked a more suggestive follow-up question.”); CECI & BRUCK, *JEOPARDY*, *supra* note 64, at 120 (discussing Cassel & Bjorklund). Fivush and Schwarzmuehler cite Laumann and Elliot as finding that “[w]hen the same questions were asked again, subjects tended to change their answers more frequently than when different questions were asked.” Fivush & Schwarzmuehler, *supra* note 63, at 568 (citing L.A. Laumann & R. Elliot, *Reporting What You Have Seen: Effects Associated with Age and Mode of Questioning on Eyewitness Reports*, 75 PERCEPTUAL & MOTOR SKILLS 799 (1992)). Although Laumann and Elliot found that children who “were asked two sets of the same leading questions were significantly more suggestible on the second set of questions than were those who were asked two opposing sets of leading questions,” Laumann & Elliot, *supra*, at 810-11, the effect was not attributable to the negative effects of repetition, but rather to the positive effects of asking “opposing sets” of questions. That is, children who were asked two sets of leading questions were not more suggestible on the second set than on the first, whereas children who were asked two sets of opposing leading questions were less suggestible on the second set than on the first. *Id.* at 811.

135. See Ceci & Friedman, *supra* note 36, at 66-67 (summarizing research by Michael Lamb and his colleagues); Lyon, *New Wave*, *supra* note 52, at 1035-36 (same).

136. Warren et al., *supra* note 50. The examples included saying to a non-responsive child, “We have to know, what did [he] do to you,” coupled with “Well what’s the truth, you haven’t told us anything,” and an interviewer saying to the child “But we were talking about [the suspect]. This is [the suspect] and what does [he] do? [pause] G’s gonna take your milkshake away.” Note that in both cases the negative consequences were paired with open-ended wh- questions.

137. See *id.*

138. Compare Lyon, *New Wave*, *supra* note 52, with Ceci & Friedman, *supra* note 36.

139. Ceci & Friedman, *supra* note 36, at 68 (“Given Lamb’s narrow definition of suggestion as a statement implying a favored reply from the child, even a single instance could call into doubt the va-

menting high rates of suggestibility inevitably tests interviews containing high proportions of extremely suggestive questions. Clearly, there is a substantial difference between rewarding every "yes" response and punishing every "no" response (the Garven study on the effects of positive and negative consequences),¹⁴⁰ and expressing disappointment four percent of the time.

E. Memory and Certainty

Children with stronger memories are less susceptible to repeated questions. Several of the studies discussed in the previous section found that children who recalled more were less likely to change their answers to repeated questions.¹⁴¹ Christine Ricci and Carol Beal repeated wh- questions (general, specific, and suppositional) and found that "none of the children who were initially accurate later answered inaccurately when the question was repeated a second time."¹⁴² The effects of certainty were anticipated by the Piagetian research on children's understanding of number: Gelman and her colleagues found that if children counted an array of objects three times rather than once (thus increasing their confidence in their answer), they were less swayed by a suggestive request to count again.¹⁴³

The effects of memory strength suggest that children will be less likely to change their responses when the repeated questions concern central details or major events rather than peripheral details or minor events.¹⁴⁴ Indeed, if the questions involve highly salient and memorable details, repetition may have no effect whatsoever. April Schwarzmüller found that four- to six-year-olds were impervious to twice repeated wh- and yes/no questions about their interactions with a "wizard," and speculated that this was because "children were active par-

lidity of an interview—and all the more so the five to ten such instances that are typical of the interviews Lamb studied.”).

140. Garven et al., *Allegations of Wrongdoing*, *supra* note 44.

141. See Candel, Merckelbach & Muris, *supra* note 123, at 66 (Bonn Test of Statement Suggestibility: shift negatively related to memory performance on free recall test); Endres, Poggenpohl & Erben, *supra* note 60, at 135 (repeating story had significant effect on reducing effects of repetition (forty-three percent to twenty-nine percent), although authors characterize effect as "weak"); Henry & Gudjonsson, *supra* note 127, at 499 (Gudjonsson Suggestibility Scale: differences between mentally retarded children and nonretarded children on shift disappeared after controlling for the effects of initial memory); Warren, Hulse-Trotter & Tubbs, *supra* note 124, at 281 (Gudjonsson Suggestibility Scale: "[C]ontrolling for initial memory (free recall) resulted in nonsignificant correlations between age and all other variables"). See also Fivush & Schwarzmüller, *supra* note 63, at 569 (discussing Warren, Hulse-Trotter & Tubbs's findings: "[t]his suggests that when the initial memory is good, children and adults are less likely to change their answers to repeated questions than when the initial memory is poor"). But see Endres, *supra* note 119, at 55 (Bonn Test of Statement Suggestibility: shift unrelated to free recall); Endres, Poggenpohl & Erben, *supra* note 60, at 135 (same); Scullin & Ceci, *supra* note 129, at 853 (Video Suggestibility Scale for Children: non-significant correlation between shift and recall).

142. Ricci & Beal, *supra* note 92, at 135.

143. R. Gelman, E. Meck & S. Merkin, *Young Children's Numerical Competence*, 1 COGNITIVE DEV. 1, 10 (1986) ("Children in the . . . group benefitted somewhat from their prior counting experience [of counting three times].").

144. Poole & White, *Question Repetition*, *supra* note 57, at 976 (noting that repetition effects probably depend on "the nature of the material to be reported (central events or details)").

ticipants” and the activities were “enjoyable and engaging for the child.”¹⁴⁵ The reader should recall, however, that even when a child’s memory of an event is quite weak, repetition does not guarantee inconsistency. In Poole and White’s follow-up study, six-year-olds asked about a minute-long interaction that had occurred two years previously were just as consistent across repeated yes/no questions as were the older children and adults.

F. Inconsistency Versus Error

The fact that children with poorer memories are more likely to change their responses when questions are repeated raises the issue of whether repetition increases error. If repetition leads children to change their wrong answers as well as their right answers, then repetition need not increase error. If children are more likely to change wrong answers than right answers, then repetition could decrease error. If repetition leads children to refuse to respond altogether or to respond “I don’t know,” then repetition could decrease accuracy without increasing error.

Evidence that repetition increases error is remarkably hard to find. As noted above, repeated open-ended questions do not decrease accuracy. With respect to yes/no questions, Poole and White could not calculate accuracy for the repeated question, “Did the man hurt Melanie?”, because subjects could reasonably answer “yes” or “no,” and they do not report accuracy for the repeated yes/no question, “Did the man ask nicely for the pen?”¹⁴⁶ Similarly, Warren and her colleagues do not report error rates across repeated yes/no questions.¹⁴⁷ A follow-up study to Warren’s, in which the authors used a similar method to examine the performance of learning disabled children in response to repeated yes/no questions, found no decrease in accuracy.¹⁴⁸ Finally, Ricci and Beal showed five-year-olds a three-minute slide show in which a camera was stolen. Error rates were unaffected by repeated questions, which included “suggestive” wh- questions (“What about the dog, what kind of dog was at the picnic?”, when there was no dog) and specific wh- questions (“What color was the ball the boys were throwing?”).¹⁴⁹

Although an “I don’t know” response is not informative, it is not inaccurate either. Summarizing Stephen Moston’s study of the effects of repeated yes/no and specific wh- questions on six- to ten-year-olds,¹⁵⁰ Ceci and Bruck emphasize that Moston found that “[t]he number of correct responses significantly de-

145. April Schwarzmüller, *The Effects of Repeated Questioning on Children’s Memory For an Event*, Poster presented at the biennial meeting of the Society for Research in Child Development, Washington, DC (Apr. 3-6, 1997) (copy on file with author).

146. Poole & White, *Question Repetition*, *supra* note 57, at 980-84.

147. Warren, Hulse-Trotter & Tubbs, *supra* note 124. Warren and her colleagues used the Gudjonsson Suggestibility Scale, *id.* at 276, which classifies a “yes” or a “no” to a “don’t know” or vice-versa as a shift. See also Gisli H. Gudjonsson, *A New Scale of Interrogative Suggestibility*, 5 PERSONALITY & INDIVIDUAL DIFFERENCES 303, 306 (1984).

148. Buhrman & Warren, *supra* note 127, at 14-15.

149. Ricci & Beal, *supra* note 92, at 131-33.

150. Moston, *supra* note 112.

clined from the first question to the second question.”¹⁵¹ Other researchers, however, have noted that the number of *incorrect* responses did not increase.¹⁵² This pattern of results is easily explained by an increase in “I don’t know” responses¹⁵³ and by a shift from one incorrect response to another.¹⁵⁴ Young children are notoriously reluctant to answer “I don’t know,” at least to yes/no questions,¹⁵⁵ but if the effects of repetition are small, then even a few “I don’t know”’s may affect the results of a study. Studying five- to eight-year-olds’ responses to repeated specific wh- and yes/no questions about a five-minute event, Amina Memon and Rita Vartoukian excluded the “I don’t know”’s from their data

151. Ceci & Bruck, *Suggestibility*, *supra* note 57, at 420.

152. See Moston, *supra* note 112, at 77 (“[R]epeated questioning did not significantly affect the mean percentage of incorrect responses given.”); Fivush & Schwarzmüller, *supra* note 63, at 569 (“[R]epeating the same questions within a single interview session does not increase [the] amount of incorrect information.”). Cf. Poole & White, *Question Repetition*, *supra* note 57, at 976 (“Moston (1987) found that repetition within a single interrogation reduced the number of correct but not the number of incorrect replies for children age 6-10.”).

153. See Fivush & Schwarzmüller, *supra* note 63, at 569 (“[R]ather than changing their response when asked the same question again, children at all ages tended to give a don’t-know response.”). Moston’s finding that repetition led children to claim ignorance appears to conflict with the claim that repetition decreases children’s apparent uncertainty. Discussing Poole and White’s initial study, Ceci and Bruck argued that

[w]hen children were asked a specific question about a detail for which they had no information (i.e., “What did the man do for a living?”), many answered with sheer speculations. Furthermore, both children and adults used fewer qualifiers with repeated questions (they omitted phrases such as “it might have been”) and consequently sounded more confident about their statements. These findings illustrate the danger of repeatedly asking specific questions: Children will often cooperate by guessing but after several repetitions, their uncertainty is no longer apparent.

Ceci & Bruck, *Children’s Testimony*, *supra* note 49, at 740. Ceci and Bruck overlooked several important findings. For example, dropping uncertainty qualifiers is problematic only if the response is incorrect. If one becomes more certain about one’s correct responses with repetition, this is a good thing. Yet when Poole and White analyzed incorrect responses, “there was no tendency for subjects as a whole to increase or decrease their use of uncertainty qualifiers. Thus, there is no evidence that question repetition encouraged subjects to speculate more or less than they did when first describing the event.” Poole & White, *Question Repetition*, *supra* note 57, at 980. In their follow-up study, Poole and White reported a decrease in uncertainty qualifiers with repetition (for both the original and the follow-up study), but included *both* correct and incorrect responses, thus making it impossible to determine whether the result is an argument against repetition. See Poole & White, *Two Years Later*, *supra* note 85, at 848. Ceci and Bruck emphasize children’s tendency to speculate about the man’s occupation. But in their initial study, Poole and White found that *adults* were more likely than children to speculate. Poole & White, *Question Repetition*, *supra* note 57, at 981 (“[A]dults speculated more frequently than did the children.”). In the two-year follow-up, the researchers found the reverse tendency: Children speculated more frequently than the adults, and the researchers suggest that this was due to the children’s weak memory for a one-minute interaction they had witnessed two years previously. Poole & White, *Two Years Later*, *supra* note 85, at 850, 852. Moreover, the researchers did indeed find decreased uncertainty when this question was repeated in the original study, but the effect was attributable to six subjects, five of whom were adults. Poole & White, *Question Repetition*, *supra* note 57, at 981-82. Hence, out of eighty-six children tested, *one* exhibited increased confidence over repeated questions. In sum, the Poole and White studies provide only mixed support for the claim that children are particularly vulnerable to appearing more certain of their responses when questions are repeated.

154. See Moston, *supra* note 112, at 77 (“If a child is guessing the answer to a question, then he or she is probably just as likely to get it wrong the first time of asking as the second time.”).

155. See Poole & White, *Tell Me Again*, *supra* note 92, at 40; Debra Ann Poole & D. Stephen Lindsay, *Children’s Eyewitness Reports After Exposure to Misinformation From Parents*, 7 J. EXPERIMENTAL PSYCHOL.: APPLIED 27, 36 (2001) (yes/no questions).

(about five percent of all responses) and found small and statistically insignificant differences in accuracy across repeated open-ended or closed-ended questions.¹⁵⁶ In sum, because children may change their wrong answers to right answers or may respond “I don’t know,” repetition may increase inconsistency without increasing error.

G. Applying the Research

Given what is now known about the effects of repetition on young children’s reports, what conclusions can be made about *Larson*? On the one hand, one can argue that the child’s statements are unreliable. The child was barely four years old when questioned. The physician’s assistant failed to ask open-ended questions. One might presume suggestiveness, based on the post-divorce timing of the allegation, and the failure of the physician’s assistant to record the interview.¹⁵⁷

Nonetheless, any expert opinion that the child would be “likely” to change her response simply because the physician’s assistant repeated the question would be unwarranted. The expert might argue that the child would be “more likely than an adult” to change her response, but would have to acknowledge that this conclusion was solely attributable to the girl’s age. No one contended that the event was ambiguous; indeed, the physician’s assistant suspected innocent touching and asked questions to test this hypothesis. No one argued that the child’s memory of interactions with her father were weak, or that the event was somehow peripheral or unimportant. No one argued that the physician’s assistant asked tag questions (“He touched your vagina, didn’t he?”) or suppositional questions (“When did your daddy touch your vagina?”). Rather, the defense contended that she asked yes/no questions, whereas the prosecution argued that she asked wh- questions. The research suggests that a four-year-old child in this girl’s situation will stand firm rather than make a false allegation in the face of a repeated question.

The child did *not* stand firm, however; she changed her story when the question was repeated. Is this not proof of inconsistency and evidence for error? This last argument raises an important question: If repetition leads to inconsistency, why should interviewers ever repeat questions? Or, to put it another way, why shouldn’t repetition that results in inconsistency always be evidence of unreliability? The answer highlights what is almost always missing from the research on suggestibility.

156. Memon & Vartoukian, *supra* note 92, at 409 (“The effect of trial on accuracy of responses to open questions fell short of statistical significance.”), 410 (closed-ended questions: “Again, the effect of trial fell short of statistical significance.”).

157. Cf. Ceci & Friedman, *supra* note 36, at 101-02 n.289 (suggesting that courts might allow expert testimony on suggestibility when the state fails to videotape a formal interview or when a case arises in the context of a particularly acrimonious divorce).

V

THE NEGLECTED ERROR: FALSE DENIALS OF ABUSE

A common-sense explanation for the child's inconsistency in *Larson* is that she was reluctant to reveal what her father had done to her. The physician's assistant testified that when she first asked the child if anyone had touched her, the child "looked down and away from me and averted her eyes."¹⁵⁸ It was only after the question was repeated that the child described how "Daddy" had touched her "between the legs" with "his bone" that "came out of his pants."¹⁵⁹ When left alone with her mother, the child claimed that she was "only kidding," but when the assistant returned she repeated the allegation and explained that "she didn't want to have her father in trouble."¹⁶⁰ In a subsequent interview she described abuse that had occurred in "her Daddy's house," but refused to name the perpetrator because "she loved her Daddy and she didn't want him to get into trouble."¹⁶¹

The possibility that abused children may be reluctant to disclose their abuse may come as no surprise. It is, however, the subject of intense debate among suggestibility researchers. Reviewing the research in a recent issue of *American Psychologist*, the house organ of the American Psychological Association, Bruck, Ceci, and Hembrooke characterized the belief that sexually abused children are reluctant to disclose abuse when directly questioned as a "stubborn urban legend."¹⁶² In their prominent book-length review of research on children's suggestibility, Ceci and Bruck characterized the belief that threats deter sexually abused children from reporting abuse as "professional 'lore.'"¹⁶³ If they are correct, and abused children are forthcoming about abuse, there is no justification for repeating questions in order to encourage a child to reveal abuse. Moreover, if a child reveals abuse only after prompting, this would be strong evidence that the allegation is false, because a truly abused child would not require such prompting.

If the researchers are wrong, however, and some children are reluctant to reveal abuse, then repeating a question may increase the number of truly abused children who will reveal their abuse when questioned. A substantial body of both observational and laboratory research demonstrates that the so-called urban legends and lore about children's reluctance to reveal abuse are true.¹⁶⁴ Children are reluctant to disclose abuse. Responding to my claims, Ceci

158. Respondent's Brief on Remand from the United States Supreme Court at 6, *State v. Larson*, 453 N.W.2d 42 (Minn. 1990) (No. C2-88-2379).

159. *State v. Larson*, 472 N.W.2d 120, 122 (Minn. 1991).

160. *Id.*

161. *State v. Larson*, 453 N.W.2d 42, 44 (Minn. 1990).

162. Bruck, Ceci & Hembrooke, *supra* note 42, at 138.

163. CECI & BRUCK, *JEOPARDY*, *supra* note 64, at 300.

164. See Thomas D. Lyon, *Scientific Support for Expert Testimony on Child Sexual Abuse Accommodation*, in *CHILD SEXUAL ABUSE* (Jon Conte ed., forthcoming 2001), available at <http://hal-law.usc.edu/users/tlyon/articles/csaas-rev.PDF> [hereinafter Lyon, *Scientific Support*]; Lyon, *New Wave*, *supra* note 52, at 1046-68.

and Friedman have acknowledged that “fear, embarrassment or loyalty may inhibit a child from disclosing abuse,”¹⁶⁵ and that “in some circumstances these considerations might make it reasonable to use suggestive questioning.”¹⁶⁶ Moreover, they have reviewed the suggestibility research and concluded that “[t]here is no serious doubt that directed questioning will often be far more effective than requests for free recall in securing disclosure of abuse.”¹⁶⁷

What effect does repetition have on reluctant children? Sadly, the effect is unknown. Suggestibility researchers have virtually ignored the extent to which suggestive techniques may overcome reluctance to report, at the same time acknowledging that any decision about the efficacy of an interviewing technique requires a consideration of the number of true accusations that the technique will elicit weighed against the number of false allegations that will occur.¹⁶⁸ Children in the studies on question repetition have little or no motive to conceal information, minimizing the likelihood that the research will reveal any positive effects of repetition.

There is a substantial body of laboratory research examining children’s reluctance to disclose minor transgressions and embarrassing events.¹⁶⁹ Although none of the studies has examined the effects of repeated questions, there is evidence that techniques criticized as “leading” increase disclosure-specific yes/no questions instead of invitations,¹⁷⁰ for example, or reassurance that it is “O.K.” to tell.¹⁷¹ In order to conduct a cost-benefit analysis of repeated questions, it is necessary to know whether repeated questions will increase the likelihood that a truly abused child will reveal abuse, as may have happened in *Larson*.

165. Ceci & Friedman, *supra* note 36, at 45.

166. *Id.* at 78.

167. *Id.* at 46. Ceci and Friedman then point out that children’s reluctance to disclose abuse can reduce the probative value of abuse allegations. Using the logic of Bayes’ theorem, they show that “all other things being equal, the less probable it is that the child would report abuse if it occurred, the less probative is a report that she does make.” *Id.* at 78. However, “all other things” are not equal; many of the factors that make truly abused children reluctant to reveal abuse also make non-abused children reluctant to allege abuse. For example, consider the fact that grade school children undergoing genital examinations are less likely to recall genital touch than other forms of touch. If “all other things” are equal, this would mean that a report of genital touch is less likely to be true than a report of non-genital touch. The same factor that suppresses true reports of genital touch (embarrassment), however, also makes gradeschool children less likely to make false claims of genital touch than less embarrassing forms of touch.

168. See CECI & BRUCK, JEOPARDY, *supra* note 64, at 2 (“[I]t 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.”).

169. See Lyon, *New Wave*, *supra* note 52; Lyon, *Scientific Support*, *supra* note 164.

170. See, e.g., J. Clare Wilson & Margaret-Ellen Pipe, *The Effects of Cues on Young Children’s Recall of Real Events*, 18 N.Z. J. PSYCHOL. 65 (1989); 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).

171. See Thomas D. Lyon & Joyce S. Dorado, *Does the Oath Matter? Motivating Maltreated Children to Tell the Truth*, Paper presented at the annual meeting of the American Psychological Society, Denver, Colo. (June 3-6, 1999) (on file with author).

VI

CONCLUSION

A careful review of the literature on repeated questions leaves one very skeptical about the value of the research for a case-by-case analysis of the reliability of a child's statements. The evidence that repetition undermines accuracy is weaker than it first appears, and the ways that repetition may increase accuracy have been ignored.

This paper is a cautionary tale for courts that find themselves confronted with claims about the reliability of children's statements. Changes in the law of hearsay make it likely that in future sexual abuse cases courts will assume more responsibility for assessing reliability. At the same time, the research on children's suggestibility continues to grow, and more frequently appears in places likely to be noticed by the courts. Moreover, there appears to be a trend toward greater acceptance of defense expert testimony on suggestibility.¹⁷² Undoubtedly, the research on suggestibility has done a lot of good. It has disabused many professionals of simplistic and over-optimistic assumptions about the possibility of false allegations. It has led to a tremendous effort to develop non-suggestive protocols for interviewing children about abuse. It has uncovered cases in which highly coercive interviewing techniques have terrorized purported child victims and potentially generated convictions of the innocent.

In the wrong hands, however, the research can also cause great harm. Adversarial battles are particularly likely to generate polarized and exaggerated claims about what science has proven. Without any evidence about the potentially beneficial effects of repetition on truly abused children, a court consulting suggestibility experts would only hear about the dangers of repeated questions. In *Larson*, the common sense interpretation of the young child's initial denial of abuse as reluctance might well be dismissed as an "urban legend" or layperson's "lore." Moreover, because of the difficulties of communicating complicated and often conflicting findings, many experts would feel compelled to simplify the literature, making it understandable to a lay audience. Simple statements about young children's unreliability when questions are repeated, however, are more likely to mislead than to educate.

172. See Lyon, *Expert Testimony*, *supra* note 68.