

# “That Never Happened”: Adults’ Discernment of Children’s True and False Memory Reports

Stephanie D. Block

University of North Carolina, Chapel Hill and University of California, Davis

Donna Shestowsky, Daisy A. Segovia, and

Gail S. Goodman

University of California, Davis

Jennifer M. Schaaf

University of North Carolina, Chapel Hill

Kristen Weede Alexander

California State University, Sacramento

Adults’ evaluations of children’s reports can determine whether legal proceedings are undertaken and whether they ultimately lead to justice. The current study involved 92 undergraduates and 35 laypersons who viewed and evaluated videotaped interviews of 3- and 5-year-olds providing true or false memory reports. The children’s reports fell into the following categories based on a 2 (event type: true vs. false)  $\times$  2 (child report: assent vs. denial) factorial design: accurate reports, false reports, accurate denials, and false denials. Results revealed that adults were generally better able to correctly judge accurate reports, accurate denials, and false reports compared with false denials: For false denials, adults were, on average, “confident” that the event had *not* occurred, even though the event had in fact been experienced. Participant age predicted performance. These findings underscore the greater difficulty adults have in evaluating young children’s false denials compared with other types of reports. Implications for law-related situations in which adults are called upon to evaluate children’s statements are discussed.

*Keywords:* children, false memory, false denials, false memory detection, child abuse

Research on children’s false memories of entire events has garnered considerable attention in recent years. Such research is propelled by the important role children’s reports can play in the legal system. A prime example stems from the crime of child

sexual abuse. This serious offense poses a special challenge for the criminal justice system because it often leaves no physical trace. Other kinds of cases that frequently involve children, such as custody or domestic violence, can also hinge considerably on children’s reports. Ideally, such reports would always be both true and complete, but the specter of children’s false reports (children’s incorrect reporting that an event occurred when in fact it did not) and, conversely, false denials (children’s incorrect denial that an event happened when it actually did occur) raises serious concerns.

Fortunately, the legal system can tolerate false reports and false denials as long as fact finders correctly assess witness accuracy. However, if adults are unable to reliably identify children’s false reports and false denials, or if adults fail to give credence to children’s accurate reports and accurate denials, serious miscarriages of justice may result. In the present study, we evaluated adults’ abilities to discern 3- and 5-year-olds’ true and false memory reports. Specifically, we examined adults’ abilities to assess children’s accurate reports, false reports, accurate denials, and false denials. To our knowledge, the present study is the first to examine adults’ abilities to differentiate children’s true from false memory reports of entire incidents.

There is longstanding concern that it can be difficult for adults to identify accurate reports made by child witnesses because of the limitations inherent in their cognitive and socioemotional development. This concern is amplified by findings that young children can be led into making false reports of entire events (e.g., Ceci & Bruck, 1993; Goodman, 1984). However, children’s denials are also of critical importance (Lyon, 1999). False denials could occur in situations in which, for example, children who were actually

---

This article was published Online First November 21, 2011.

Stephanie D. Block, Department of Psychology, University of California, Davis and Center for Developmental Science and The Injury Prevention Research Center, University of North Carolina, Chapel Hill; Daisy A. Segovia, and Gail S. Goodman, Department of Psychology, University of California, Davis; Donna Shestowsky, School of Law, University of California, Davis; Jennifer M. Schaaf, Department of Psychology, University of North Carolina, Chapel Hill; Kristen Weede Alexander, Department of Child Development, California State University, Sacramento.

Stephanie D. Block is now at the Department of Psychology, University of Massachusetts at Lowell.

Daisy Segovia is now at the John Jay College of Criminal Justice, City University of New York.

We thank Amanda Chavez, Diane Russell, Samantha Maunder, and Sascha Zubryd for their research assistance. We also thank participants of the Law and Economics Seminar at the University of Texas, Austin, School of Law, for comments on a draft of this article. Special gratitude is extended to Kevin Grimm for statistical consultation and Mark Hoffman for video editing. This research was funded in part by the American Society for Trial Consultants (Stephanie Block) and the Office of Research at the University of California, Davis (Donna Shestowsky).

Correspondence concerning this article should be addressed to Donna Shestowsky, University of California, Davis, School of Law, 400 Mrak Hall Drive, Davis, CA 95616. E-mail: dshest@ucdavis.edu or Gail S. Goodman, Department of Psychology, University of California, 1 Shields Avenue, Davis, CA 95616. E-mail: ggoodman@ucdavis.edu

abused deny the abuse, and accurate denials could occur in cases in which a third party suspects wrongdoing that did not in fact take place. For justice to prevail, adults cannot treat any of these four types of child statements with misplaced credence or unwarranted skepticism.

For our study, adults evaluated interviews of children who were suggestively questioned about true and false events but not instructed to intentionally lie. Instead, we used a variant of a false memory paradigm (e.g., Loftus & Pickrell, 1995). In response to the interviews, children spontaneously answered with true reports, true denials, false reports, or false denials. Thus, the false reports and false denials used as stimuli did not necessarily reflect deliberate, premeditated deception, although that is one possibility. Children might falsely assent or falsely deny for a variety of reasons, including succumbing to demand characteristics during an interview (e.g., Bjorklund et al., 2000; Cassel, Roebbers, & Bjorklund, 1996), using their imaginations (Reyna, Holliday, & Marche, 2002), being deficient in certain social-cognitive realms (e.g., incomplete understanding of honesty, Bussey, 1992; limited theory of mind or inadequate inhibition, Talwar & Lee, 2008), having been coached or coerced into lying (e.g., Ackil & Zaragoza, 1998; Lyon, Malloy, Quas, & Talwar, 2008), trying to conform to social expectations of politeness (e.g., Talwar, Murphy, & Lee, 2007), or acquiescing to persistent questioning (e.g., Schaaf, Alexander, & Goodman, 2008). Children who falsely deny an event also might do so to avoid talking about an unpleasant or embarrassing experience (Bruck, Ceci, & Hembrooke, 1998), evade punishment (e.g., Polak & Harris, 1999; Talwar, Gordon, & Lee, 2007), or reduce the chances of getting someone else in trouble (e.g., Goodman-Brown, Edelstein, Goodman, Jones, & Gordon, 2003).

In the current study, adults evaluated true and false memory reports made by preschoolers. Young children were of special interest because they are particularly vulnerable to the aforementioned developmental limitations and social forces, including false memory reports. Despite such limitations, research reveals that children make great strides in cognitive and social development during the preschool years, in concert with increased resistance to suggestion (Harris, Goodman, Augusti, Chae, & Alley, 2009). In addition, children's social awareness and understanding of truth and lies improve substantially during this developmental period (e.g., Lee & Cameron, 2000). For such reasons, researchers have paid particular attention to the accuracy and suggestibility of preschoolers' eyewitness memory (e.g., Goodman & Reed, 1986; Poole & Lindsay, 1998). Yet, there is surprisingly little research on adults' abilities to discern preschoolers' true memory from false memory accounts of entire events.

### Adults' Discernment of True and False Statements

A growing body of research exists on adults' abilities to discern children's accurate and inaccurate responses when answering questions about details or actions within experienced (true) events (e.g., Ball & O'Callghan, 2001; Leippe, Manion, & Romanczyk, 1993). Typically, this research shows that adults have difficulty identifying inaccurate reports. For example, Leichtman and Ceci (1995a, 1995b) asked researchers and clinicians to view videotapes of three 3- to 5-year-olds nonrandomly selected (of 39 children) to give coherent, confident accounts (some accurate and some false) of "Sam Stone's" brief visit to their classroom. The children had

been subjected to many weeks of repeated stereotyping and suggestive interviewing. Then, at the start of conference talks, researchers and clinicians in the audience were asked to view the videotapes and quickly rate their confidence in whether or not the event had occurred and the children's credibility. Under these circumstances, the professionals were no better than chance at detecting true reports (e.g., a short but accurate account) from false reports (e.g., a detailed but false account). Leichtman and Ceci concluded that it is generally difficult to detect young children's false memories (see also Ceci & Huffman, 1997). In more recent research, Laimon and Poole (2008) presented adults with videotaped interviews of three nonrandomly selected 4- to 5-year-olds (of 114 children) being asked about falsely suggested information regarding a true event, a visit by "Mr. Science" to their classroom. Most (75–89%) of the adults incorrectly judged children's erroneous reports as accurate when the children provided details via free recall or after assenting to a yes–no question. In contrast, when children denied the yes–no question and then, when pressed further, provided inaccurate details, only 8% to 30% of the adults deemed false reports as accurate. However, because only children's false reports about a true event were included, it is unclear whether these findings generalize across event types or across children generally. Overall, in studies on children's false reports of details of true events, adults often underestimate children's degree of memory error. However, especially with young children, adults at times overestimate children's error rates (Leippe et al., 1993).

Although we know of no previous studies on adults' discernment of children's true and false memories of entire events, research on adults' discernment of *adults'* true and false memory reports is more established (e.g., Ost, Vrij, Costall, & Bull, 2002; Schooler, Gerhard, & Loftus, 1986). For example, Qin, Ogle, and Goodman (2008) examined undergraduates' judgments of adults' true and false childhood memories. The observers distinguished between true and false memories at an above-chance level, but error rates were high. True memories were rated higher than false memories in regard to several memory characteristics (e.g., plausibility). Event plausibility is thus a potentially important factor in identifying false memory reports. Taken across studies, observers show some ability to detect adults' true from false memories, although the general level of accuracy is unimpressive. Nevertheless, some characteristics could potentially be used to reliably differentiate adults'—and perhaps children's—true and false memories. False memory reports tend to be shorter, contain more verbal hedges (e.g., "I'm not sure but . . ."), less perceptual detail, and fewer cognitive operations (e.g., "I remember thinking . . .") than true memory reports (e.g., Heaps & Nash, 2001; Hyman & Pentland, 1996; Loftus & Pickrell, 1995). These studies suggest that it may be possible for observers to discern true from false memories.

### Perceptions of Children

When evaluating children's statements, adults are likely to consider not only statement content, but also how they perceive the children making the statements (Goodman et al., 2006; Strömwall & Granhag, 2005). Orcutt, Goodman, Tobey, Batterman-Faunce, and Thomas (2001) and Brigham (1998) found that ratings of children's accuracy and believability were the strongest predictors of guilt ratings in mock trials involving child witnesses. Castelli,

Goodman, and Ghetti (2005) reported that when young children are interviewed in a highly leading manner, they are viewed as miscommunicating what occurred (see also Landström, Granhag, & Hartwig, 2007). To further elucidate contributions of child witness-related characteristics to adults' discernment, we examined whether, compared with children who provide accurate reports or accurate denials, children who make false reports or who falsely deny events are perceived as less accurate and less believable and as miscommunicating more.

### Individual Differences

Research attests to reliable individual differences in discernment of true versus false reports (e.g., Edelstein, Luten, Ekman, & Goodman, 2006). For example, older observer age and greater experience with children can affect adults' abilities to discern children's truth-telling from falsehood (e.g., Leippe et al., 1993; Talwar & Lee, 2002). Observer gender is an individual-difference factor that may also affect discernment accuracy. Although overall truth biases may exist in judging children's statements (Strömwall & Granhag, 2005; Strömwall, Granhag, & Landström, 2007), Edelstein et al. (2006) found female versus male observers to be more biased toward making “truth” judgments, but they did not find gender differences in accuracy. Similarly, mock jury studies suggest that females are particularly likely to believe child witnesses (and convict defendants) when children testify as victims in sexual abuse cases (Bottoms, Golding, Stevenson, Wiley, & Yozwiak, 2006; Castelli et al., 2005; but see Leach, Talwar, Lee, Bala, & Lindsay, 2004).

Finally, individual differences in preexisting views and attitudes about children may affect adults' detection of children's true and false statements. In Norway, Magnussen et al. (2006) discovered that the public generally believed that children had better mnemonic capacity than adults. In their study of discernment of adults' true and false memories, Qin et al. (2008) reported that attitudes toward memory (i.e., belief in the special status of trauma memory) predicted lower  $d'$  scores and more lenient biases to judge memories as true. Thus, it is possible that general attitudes toward child witnesses (e.g., children's cognitive abilities as witnesses, children's credibility) would predict discernment of children's true versus false memory reports as well.

### The Present Study

For the present study, undergraduates and laypersons viewed videotaped interviews of 3- and 5-year-olds who provided accurate reports, false reports, accurate denials, or false denials. The children, videotaped for a separate study of true and false memory (Schaaf et al., 2008), were interviewed with repeated, strongly leading questions within a single session about events known (based on parental report) to have been experienced or not by the children. The questions were meant to be analogous to those that might be asked in a forensic context when the interviewer has a preconceived idea of what occurred and is using repeated leading questions to elicit a report of the incident. A similar series of questions was applied to all event types and for both true and false events. Interviewers did not know the veracity of the events.

For the present research, children were randomly selected from the previous study to fill four experimental cells created by a 2

(event: true vs. false event)  $\times$  2 (child's report: assent vs. denial) factorial design, while trying to achieve an equal balance between children's age and gender groups and children's reporting of positive and negative events. We also ensured that the children's responses met the criteria for our experimental design. In particular, no more than one response per child to the six questions could depart from our categorization (as a true or false report or true or false denial). The criteria permitted our categorization to be maintained while still accommodating some inconsistency in the children's responses, as is often characteristic of young children (Fivush, 1993). Videotapes of the children answering the repeated questions were presented to undergraduates and adult laypersons who were asked to make judgments about the children and their reports. One such judgment—and the key one of interest—was how confident the adults were that the children had actually experienced the event.

### Hypotheses

We advanced several hypotheses. First, we predicted that observers would be more accurate—as reflected in their confidence-that-the-event-occurred ratings—in identifying true memory reports (accurate reports, accurate denials) than false memory reports (false reports, false denials) because of a “truth” bias (Chahal & Cassidy, 1995; Landström et al., 2007; Strömwall & Granhag, 2005). This “bias” stems from the fact that adults are often more likely to make a judgment of truth than of deception, thereby potentially leading to errors in false report and false denial detection. Second, we hypothesized that to the extent participants could discern true from false memory reports, accuracy and believability would be rated higher for children who provided accurate reports and accurate denials, whereas miscommunication would be rated higher for children who provided false reports and false denials (Landström et al., 2007; Strömwall & Granhag, 2005; Vrij, Edward, & Bull, 2001). Third, consistent with previous research (Edelstein et al., 2006), we expected women to be more biased than men in identifying children's statements as true. Fourth, we hypothesized that older age or greater experience with children would be associated with more accurate judgments of children's statements. Fifth, we anticipated that individuals who give child witnesses in general more favorable rather than less favorable ratings regarding competency, ability as witnesses, credibility, and cognitive capability would be especially likely to judge children's statements as true.

### Method

#### Participants

A total of 92 undergraduates ( $M = 20.98$  years; range 18 to 38 years; 58 women) participated in exchange for course credit. In addition, 35 laypersons ( $M = 37.03$  years; range 19 to 79 years; 26 women) from the community were recruited via posted fliers and local newspapers and were paid for their participation. The sample was ethnically diverse (undergraduates: 23% Caucasian, 2% African American, 12% Hispanic, 51% Asian, 12% other ethnicities; laypersons: 66% Caucasian, 3% African American, 9% Hispanic, 14% Asian, 9% other ethnicities). Although the limited number of laypersons constrained our statistical power to detect “participant

type” differences, inclusion of adult community members was important to capture variations in age, experience with children, and attitudes toward child witnesses that could influence adults’ judgments of children’s statements.

## Materials

**Demographics.** A standard demographic questionnaire (e.g., concerning education level, marital status, parenthood, and race/ethnicity) was included.

**Videotaped stimuli.** Videoclips of interviews from a study of children’s memory for true and false events were used (Schaaf et al., 2008). In that study, parental reports were gathered on three separate occasions to confirm the accuracy of true and false events. First, prior to the laboratory visit, parents were interviewed in detail over the phone about experiences their child may have had. For example, if a child had visited Disneyland during the past year, the parent was asked about the child’s favorite ride and who was there during the event. Following this initial parent contact, the child’s interview was created, tailored to his or her experiences (but standardized in general question content and format, with most questions randomly ordered; see Appendix). The interview included six questions about each event: four open-ended and two yes–no questions. Second, on arrival at the laboratory, parents were asked to carefully review the interview for accuracy of true events and to affirm that the false events and all related details did not occur. Third, parents watched the child being interviewed, via closed-circuit TV. Parents provided additional information about children’s accuracy at that time. There is precedent for relying on parental reports to determine children’s overall accuracy of true and false events as well as specific details of the events (e.g., Ceci, Loftus, Leichtman, & Bruck, 1994; Ghetti & Alexander, 2004).

Twenty-four videoclips were equally divided across the 2 (event type)  $\times$  2 (child report) factorial. In addition, one practice clip (in which a child provided an accurate report) was included. Thus, in all, participants were exposed to reports by 25 different children. In each videoclip, the child and interviewer (some male, some female) were seated side by side at a table, with both in view. The interviews included a straightforward series of questions about each event, with all questions in the series asked even if the child denied the event’s occurrence. For true events, the child either correctly assented to the event (accurate report) or incorrectly denied that the event happened (false denial). For false events, the child either incorrectly assented to the event (false report) or correctly denied that the event happened (accurate denial).<sup>1</sup> Equal numbers of 3- and 5-year-olds were chosen for each event and report type, with gender roughly equated (e.g., child denied event: six girls, six boys; child affirmed event: five girls, eight boys, including the practice child). The ethnicities of the children in the videos were 17 White, one Latino, two Asian American, two other, three unknown. Each child was asked about an event that fell into one of the following categories: punishment (e.g., getting in trouble for throwing a rock through a window), gift (e.g., getting a big stuffed yellow giraffe), embarrassment (e.g., having another child pull down the child’s pants in front of a store), or trip (e.g., going to Disneyland). The events were selected to represent different emotional valences of experiences in the lives of children. In addition, the punishment and embarrassment events were selected because children are often reluctant to discuss such events because

of, for example, unpleasant associations, as well as possible negative consequences for making such revelations. In these respects, the events shared some similarity to events in forensic situations. The true events were real-life events (i.e., events that actually occurred) selected by the parents within these four categories. False events were selected from the true events parents described during the pilot phase of the original study but were false for the child being interviewed (see Schaaf et al., 2008).

Two versions of the videoclips were created to counter possible order effects. Each version started with the same practice clip. Each clip thereafter was randomly assigned a position 1 through 24, with the exception that no more than three child assents or denials and no more than three of the same event category (e.g., gift, trip) could occur in a row.

**Children’s Statement Evaluation Questionnaire.** This questionnaire, designed for the current study, assessed observers’ confidence that each child actually experienced the event of interest (rated on a 6-point scale, with 1 = *extremely confident s/he did not* and 6 = *extremely confident s/he did*), which was our primary dependent measure. This questionnaire also included 5-point scales (1 = *not at all* and 5 = *extremely*) for ratings of each child’s characteristics, such as honesty, believability, and ability to recall the event, and one such scale to rate the plausibility of the event. A separate rating sheet containing these questions was included for each child’s interview. Each sheet started with the confidence question, followed by the other questions appearing in one of two random orders.

**Experience with children.** To measure the amount of experience participants had with children, we included the following question: “How much direct experience have you had with children?” (1 = *a little* to 6 = *a lot*).

**Attitudes Toward Child Witnesses (Wrightsmen, Batson, & Edkins, 2004).** On this questionnaire, a 7-point Likert scale (1 = *strongly disagree* to 7 = *strongly agree*) indicates degree of agreement with each of 21 statements about child witnesses. Factor analysis by the developers indicated four attitude subscales: Children’s Competence, Views of Child Witnesses at Trial, Children’s Credibility, and Views of Children’s Abilities as Witnesses.

## Procedure

Participants, tested in groups that ranged from two to 15 ( $M = 6.05$ ), were seated in front of a large TV monitor. After providing consent, participants completed the demographic questionnaire. The Children’s Statement Evaluation Questionnaire was then distributed. Participants were instructed to take the task “seriously and thoughtfully—just like you would if you were making deci-

<sup>1</sup> Children at times said, “I don’t know.” Across the six questions asked, this answer was provided two times or less per child, with one exception, which involved a 3-year-old in the false denial condition. The child falsely denied the event, but also provided five “I don’t know” responses. We included this child to best capture the variability in young children’s false denial reports. It was furthermore of interest to determine whether the number of words provided differed in our four main conditions (e.g., false denial, true report). We thus counted the words spoken in each interview shown. There were no significant differences in interviewer word count or child word count in relation to event type, child report, child age, or child gender.

sions as a juror for a real case.” Participants were informed that they would watch 25 short videoclips of 3- and 5-year-olds and that in each clip, a single child would respond to a series of questions about a certain event. It was emphasized that the adults interviewing the children did not know whether the children were accurate and that the instructions observed were “the only instructions given to the children, [and] no child was instructed to lie.” In actual forensic interviews or trials, it is standard for children to be instructed by authorities not to lie.

After the practice clip, the remaining 24 interviews were presented. The experimenter paused after each interview to instruct participants to answer the questions in relation to the child just viewed. The experiment continued once all participants were ready to proceed. After viewing the videotaped interviews, participants completed the question about direct experience with children and then the Attitudes Toward Child Witnesses questionnaire (Wrightsmann et al., 2004). Finally, all participants were debriefed and thanked.

**Results**

Unless otherwise indicated, we report a series of 2 (participant type: undergraduate vs. layperson) × 2 (participant gender) × 2 (event type: true vs. false) × 2 (child report: assent vs. denial) analyses of variance (ANOVAs), with event type and child report as the only within-subject factors. The dependent measures were created as mean ratings across interviews in each cell of the 2 (event type) × 2 (child report) factorial design, unless specified otherwise. All significant effects are reported. Moreover, 95% confidence intervals are presented.

Preliminary analyses indicated that, contrary to prediction, self-reported direct experience with children was not a significant predictor of the dependent measures. Thus, it is not considered further. Because differences in ethnic composition of the undergraduate and layperson samples created a potential confound, all ANOVAs were conducted again with ethnicity (majority = Caucasian, non-Hispanic vs. minority = all others) as a covariate. The

majority versus minority distinction was justified based on the composition of the participant-type groups, as well as on United States demographics in which Caucasians are the majority (including on juries; e.g., Diamond & Dolan, 2010). Effects are reported only if still significant with ethnicity covaried.

**Factor Analysis**

To reduce the number of variables from the Children’s Statement Evaluation Questionnaire, we performed a principal components analysis, with oblimin rotation. Two factors were extracted (eigenvalues >1). These factors accounted for 43.72% and 25.84% of the variance, respectively. The first factor represented Perceived Accuracy and Believability (coefficient  $\alpha = .87$ ) and included ratings of the accuracy of the child’s memory, the child’s honesty, the child’s ability to recall, and the child’s believability (factor loadings  $\geq .76$ ). The second factor reflected Miscommunication (coefficient  $\alpha = .68$ ) and included ratings of how well the child misunderstood the interviewer’s question, how hard the child was to understand and how much the child reported unrelated information (factor loadings  $\geq .70$ ).

**Confidence of Event Occurrence Ratings**

The main dependent variable of interest concerned adults’ ratings of their confidence that each event actually occurred. The full model ANOVA revealed significant main effects of event type (true,  $M = 3.01$ ,  $SD = 0.60$ , CI [2.88, 3.14]; false,  $M = 2.31$ ,  $SD = 0.63$ , CI [2.17, 2.44]),  $F(1, 123) = 245.50$ ,  $p < .001$ ,  $\eta_p^2 = .67$ , CI [.55, .76], and child report (assent,  $M = 3.34$ ,  $SD = 0.66$ , CI [3.20, 3.49]; denial,  $M = 1.98$ ,  $SD = 0.62$ , CI [1.86, 2.13]),  $F(1, 122) = 488.10$ ,  $p < .001$ ,  $\eta_p^2 = .80$ , CI [.70, .86], which were subsumed by a significant Event Type × Child Report interaction,  $F(1, 122) = 50.01$ ,  $p < .001$ ,  $\eta_p^2 = .29$ , CI [.17, .41]. As Figure 1 reveals, when events had actually been experienced, and children affirmed the events’ occurrence (i.e., accurate reports), adults were the most confident that the events had occurred. Moreover, adults

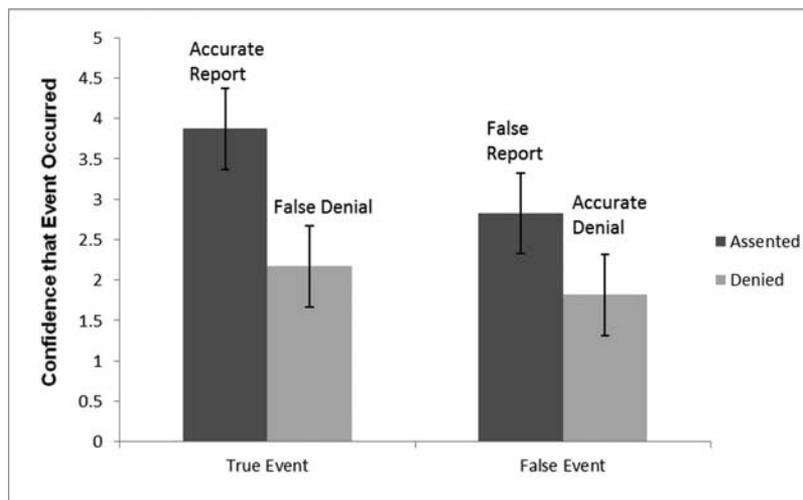


Figure 1. Participants’ mean confidence that child experienced the event (rated on a 6-point scale, with 1 = extremely confident s/he did not and 6 = extremely confident s/he did) by event type and child report.

were confident that the events had not occurred when children accurately denied the events' occurrence.

Fortunately, adults were also hesitant to believe that the events occurred when children made false reports. The biggest error, however, concerned false denials: When children falsely denied that the events had occurred, adults overbelieved the children. Specifically, simple effects analyses indicated that observers were significantly more confident that the event happened when children made accurate reports compared with when they made false denials,  $F(1, 126) = 828.02, p < .001$ , or false reports,  $F(1, 126) = 345.99, p < .001$ . Observers were more confident that the event had not happened when children made accurate denials than when they made false reports,  $F(1, 126) = 241.37, p < .001$ , or false denials,  $F(1, 126) = 58.05, p < .001$ . Of particular note, the mean for false denials represented a judgment of "confident the event had *not* occurred," even though the event had in fact been experienced. Thus, the mean was far from an accurate judgment about the children's experiences. The mean for false reports was more accurate, representing "somewhat confident the event had *not* occurred," which was closer to a correct judgment because, in fact, the children had not experienced those events. There was also a significant main effect of gender,  $F(1, 123) = 4.86, p = .03, \eta_p^2 = .04, CI [.01, .12]$ : Men ( $M = 2.81, SD = 0.71, CI [2.58, 2.99]$ ) had higher ratings than women ( $M = 2.53, SD = 0.62, CI [2.40, 2.67]$ ). The main effect of participant type was not significant,  $F(1, 123) = 3.44, ns$ .

There were too few children per age group to include child age as an additional factor in our analyses. However, when the observers' confidence ratings concerning whether the event occurred were examined across child age, the trends were identical to those reported with one exception: Observers were (incorrectly) more confident that the event had not happened when 5-year-olds ( $M = 1.57, SD = 0.62, CI [1.46, 1.68]$ ) as opposed to 3-year-olds ( $M = 2.76, SD = 1.02, CI [2.58, 2.94]$ ) made false denials. In fact, observers were the most confident of all that the event had not occurred when 5-year-olds made false denials (e.g., compared with 3- or 5-year-olds' accurate reports, false reports, or accurate denials).

### Perceptions of Children

We were also interested in adults' perceptions of the children as reflected on the Children's Statement Evaluation Questionnaire. When the accuracy/believability scores were entered into an ANOVA, significant main effects of event type (true,  $M = 3.33, SD = 0.44, CI [3.26, 3.40]$ ; false,  $M = 3.07, SD = 0.50, CI [2.98, 3.16]$ ),  $F(1, 123) = 31.87, p < .001, \eta_p^2 = .21, CI [.10, .32]$ , and child report (assent,  $M = 2.70, SD = 0.41, CI [2.62, 2.78]$ ; denial,  $M = 3.81, SD = 0.52, CI [3.69, 3.90]$ ),  $F(1, 123) = 270.71, p < .001, \eta_p^2 = .69, CI [0.56, 0.80]$ , were qualified by a significant Event Type  $\times$  Child Report interaction,  $F(1, 123) = 309.48, p < .001, \eta_p^2 = .76, CI [.68, .82]$ . Simple effects analyses indicated that children who provided accurate reports ( $M = 3.02, SD = 0.36, CI [2.96, 3.08]$ ) or accurate denials ( $M = 4.04, SD = 0.53, CI [3.95, 4.13]$ ) were rated as more accurate/believable than children who gave false reports ( $M = 2.24, SD = 0.46, CI [2.16, 2.32]$ ). In addition, children who made accurate denials were viewed as more accurate/believable than children who made false denials ( $M = 3.57, SD = 0.51, CI [3.48, 3.66]$ ), and in turn, children who made

false denials were viewed as more accurate/believable than children who made accurate reports.

ANOVA was also performed on observers' ratings of whether each child miscommunicated. Significant main effects of event type (true,  $M = 1.95, SD = 0.33, CI [1.89, 2.01]$ ; false,  $M = 2.24, SD = 0.38, CI [2.16, 2.31]$ ),  $F(1, 123) = 142.90, p < .001, \eta_p^2 = .54, CI [.45, .65]$ , and child report (assent,  $M = 2.45, SD = 0.42, CI [2.37, 2.53]$ ; denial,  $M = 1.74, SD = 0.30, CI [1.68, 1.79]$ ),  $F(1, 123) = 483.60, p < .001, \eta_p^2 = .80, CI [.71, .85]$ , were subsumed by a significant Event Type  $\times$  Child Report interaction,  $F(1, 123) = 42.13, p < .001, \eta_p^2 = .26, CI [.13, .40]$ . Children who provided false reports ( $M = 2.67, SD = 0.48, CI [2.57, 2.77]$ ) were rated as being more likely to have miscommunicated compared with children who gave accurate reports ( $M = 2.23, SD = 0.35, CI [2.15, 2.30]$ ) or accurate denials ( $M = 1.80, SD = 0.28, CI [1.74, 1.86]$ ). Surprisingly, children who made false denials ( $M = 1.67, SD = 0.32, CI [1.60, 1.74]$ ) were viewed as having miscommunicated less compared with children who made accurate reports or accurate denials. Overall, for accuracy/believability and miscommunication ratings, the results indicate a tendency to believe true and false denials and to be more skeptical of true and false reports.<sup>2</sup>

### Event Plausibility Ratings

It was also important to examine the adults' ratings of the plausibility of each event. Specifically, participants were asked how plausible (realistic) it was for 3- to 5-year-olds to experience the event in question (e.g., throw a rock through a window; 1 = *not at all plausible* and 5 = *extremely plausible*). The analysis failed to reveal significant effects with ethnicity covaried: accurate report ( $M = 3.97, SD = 0.53, CI [3.88, 4.06]$ ), accurate denial ( $M = 3.71, SD = 0.66, CI [3.61, 3.82]$ ), false report ( $M = 3.92, SD = 0.57, CI [3.82, 4.02]$ ), false denial ( $M = 3.91, SD = 0.67, CI [3.79, 4.03]$ ). Moreover, when the plausibility ratings were covaried (as a within-subject factor) from the confidence of event occurrence ratings per the 2 (event type)  $\times$  2 (child report) design using hierarchical linear modeling, the Event Type  $\times$  Child Report interaction was still significant,  $F(1, 122) = 818.97, p < .001$ , and the pattern of the means was identical to that described above. Also of interest, correlational analyses revealed that plausibility judgments significantly predicted confidence of event occurrence as well as accuracy and believability ratings for true and false reports ( $r_s = .20$  to  $.27, p_s < .05$ ), but not for true and false denials.

### Signal Detection Analyses

To examine the accuracy of ratings while controlling for response bias, signal detection theory (SDT) analyses were per-

<sup>2</sup> Note that the questionnaire also contained various additional questions (e.g., child was hesitant, child's report was based on fantasy, child was trying to deceive the interviewer, and child seemed intelligent) and accompanying 5-point scales (1 = *not at all*, 5 = *extremely*). The pattern of results for responses to these questions closely mimicked that for the accuracy/believability and miscommunication measures (e.g., significant main effects of event type and child report subsumed by significant Event Type  $\times$  Child Report interactions). Details of these analyses are available from the lead authors.

formed (Green & Swets, 1966; MacMillan & Creelman, 1991). For  $d'$ , values less than 0 indicate a lack of sensitivity. Lower values of  $\beta$  indicate a more liberal response criterion. Because  $d'$  and  $\beta$  are not defined for hits or false alarms of 0 or 1, the data were transformed by  $1/(2N)$  for 0 and  $1 - 1/(2N)$  for 1. Two  $d'$  and  $\beta$  values, respectively, were created: accurate reports versus false reports ( $d'$  report,  $\beta$  report) and accurate denials versus false denials ( $d'$  denial,  $\beta$  denial). The  $d'$  and  $\beta$  scores were entered into separate 2 (participant type)  $\times$  2 (participant gender)  $\times$  2 (SDT index: report vs. denial) mixed ANOVAs, with the latter factor varying within subjects. A significant main effect of SDT index emerged in the analysis of  $d'$  report versus denial,  $F(1, 122) = 38.43, p < .001, \eta_p^2 = .24, CI [.13, .37]$ . Participants'  $d'$  report scores ( $M = 0.99, SD = 0.81, CI [0.85, 1.13]$ ) were significantly higher than their  $d'$  denial scores ( $M = 0.25, SD = 0.59, CI [0.15, 0.35]$ ), reflecting greater difficulty discerning denials compared with reports. When  $\beta$  report and  $\beta$  denial were compared in an ANOVA analogous to that mentioned earlier, there were no significant effects involving gender or participant type, but the main effect of SDT index was significant,  $F(1, 123) = 21.57, p < .001, \eta_p^2 = .15, CI [.07, .25]$ ;  $\beta$  reports ( $M = 4.12, SD = 3.10, CI [3.58, 4.66]$ ) versus  $\beta$  denials ( $M = 1.96, SD = 2.02, CI [1.61, 2.31]$ ). Thus, participants set a more conservative criterion (were more “nay saying”) for reports than denials.

Regression analyses were conducted to predict  $d'$  and  $\beta$  scores. Age, gender, race/ethnicity (Caucasian = 1, all others = 0), Views of Child Witnesses at Trial subscale score, and Views of Children's Abilities as Witnesses subscale scores were entered as predictors (participant type was too highly correlated with age to be included and was not a significant predictor when it replaced age in the regressions). Only two of the Attitudes Toward Child Witnesses scales could be included to avoid problems of multicollinearity. For  $d'$  denial and  $\beta$  denial, the overall models were significant,  $F_s(5, 124-125) \geq 3.32, p_s < .01, d'$  denial  $R^2 = .12, CI [.02, .23]$ , and  $\beta$  denial  $R^2 = .12, CI [.02, .22]$ . Age was the only significant predictor:  $d'$  denial  $\beta = .31$  (unstandardized  $\beta = .015, CI [.01, .03]$ ), and  $\beta$  denial  $\beta = .32$  (unstandardized  $\beta = .05, CI [.02, .09]$ ;  $p_s < .05$ ). Being older predicted greater accuracy and a more conservative response bias regarding children's denials. The remaining regressions were not significant.

## Discussion

For the justice system to effectively respond to crimes or civil injustices involving children as victims or key eyewitnesses, adults within the legal system face the difficult task of evaluating children's reports. We sought to shed light on adults' confidence and accuracy in making such evaluations and on factors that adults rely on to make their judgments. We did so by expanding previous research in important ways. For example, we included not only accurate and false reports but accurate and false denials. Furthermore, we investigated evaluations of children's reports of whole experiences that were either true or false and about which children were instructed not to lie. In doing so, a major strength of our approach was to include interviews from a variety of children about a variety of events. Such “stimulus sampling” (that is, inclusion of different children recalling different events) increases faith that our findings reflect the study's experimental conditions

rather than adults' impressions of a particular child (Wells & Windschitl, 1999).

## Adults' Discernment: Confidence of Event Occurrence Judgments

We hypothesized that adults would be better able to detect true accounts than false ones because of a “truth bias,” as identified by previous researchers (e.g., Strömwall & Granhag, 2005). If our participants had accurately discerned true from false reports, they should have been confident that the accurate-report events had occurred and that the false-report events had not occurred. On average, adults were hesitant to express confidence that any of the events had happened. However, to their credit and consistent with our hypothesis, the adults were most confident that the events had occurred when the children gave an accurate report. These reports were rated, on average, as “somewhat confident the event occurred.” Moreover, when children provided accurate denials, adults' mean ratings were particularly appropriate: Adults were “confident the event had not occurred.” This pattern of results could result from a truth bias.

However, unexpectedly, false reports were also rated with relative aptitude; on average, adults were “somewhat confident” that the children had *not* experienced the false events, even though the children claimed to have experienced the events in question. This was particularly surprising given that past researchers concluded that adults cannot accurately identify children's false reports (e.g., Leichtman & Ceci, 1995a).

In contrast, it was ratings of false denials that were particularly inaccurate, with adults “confident the event did not occur” when, in reality, the events had occurred. Our results suggest that adults are likely to be particularly fooled by false denials (Lyon, 1999). That is, they are likely to have a “denial bias” through which denials are overly believed.

In that regard, it is also important to consider the pattern of our signal detection findings. These analyses necessarily collapsed across two of the four cells of the study. The signal detection analyses revealed that  $d'$  report was significantly higher than  $d'$  denial, confirming that participants had greater difficulty discerning true and false denials than true and false reports. Furthermore, participants set a more conservative response criterion (were more “nay saying”) for reports than denials. Thus, the adults truly had less sensitivity in assessing denials than reports and also were biased toward believing denials and against believing reports. To our knowledge, ours is the first study to demonstrate these effects.

To the extent that these results might generalize to criminal cases, the findings imply that adults would be unlikely to believe that an event such as abuse had happened if the child falsely denied it (absent other evidence). However, our findings also support the idea that adults may at times be led astray by false reports (e.g., Laimon & Poole, 2008): In our study, adults were more prone to believe an event actually occurred when children falsely reported the event than when children accurately denied it.

## Adults' Perceptions of Children

Adults' ratings of children's accuracy and believability and of their possible miscommunication provide another window on adults' discernment skills. We predicted that children who pro-

vided accurate reports and accurate denials would be rated most positively. Instead, there was an overall tendency for participants to rate children positively if the children denied the events. In this regard, a “denial” bias was more apparent than a “truth” bias (Strömwall & Granhag, 2005; Strömwall et al., 2007). Nevertheless, the findings partly supported our hypothesis. In general, when children reported on true events and when they denied events, adults rated the children in a more positive way (i.e., higher in accuracy/believability, having miscommunicated less) compared with when the children made false reports. The related Event Type  $\times$  Child Report interaction effects were intriguing. The pattern of our findings indicated that children who made false reports were viewed the most negatively (they were judged the least accurate/believable and the most likely to have miscommunicated) compared with children who told the truth (i.e., made accurate reports or accurate denials) or made false denials—results that as a whole cannot be explained by a difference in the plausibility of the events. This pattern for false reports is encouraging from a justice perspective because it suggests that adults intuit some of the core problems inherent in false reports (even though our results suggest that adults can still be misled by false reports).

### Event Plausibility

Although we did not advance specific predictions about event plausibility, past research suggests its potential importance in decisions about false memory reports (e.g., Qin et al., 2008). In the present study, regarding accurate and false reports, observers who viewed an event as more plausible were significantly more likely to express confidence that the event occurred and to view the child as accurate and believable. Thus, when the child claimed that an event took place (e.g., throwing a rock through a window), adults appeared to rely at least somewhat on event plausibility in deciding to believe the child. However, at least for the events we studied, when a child denied the event, observers did not seem to use plausibility as a judgment guide. It makes sense that, although an event may be plausible (e.g., going to Disneyland), the child may or may not have experienced it. In an actual legal case, however, plausibility might still be an important factor in evaluating children’s denials, and judgments of event plausibility would likely vary depending on other evidence presented in the case (e.g., if a child tested positive for a sexually transmitted disease, how plausible is it that the child was not abused?).

### Individual Differences

We predicted that gender, participant age, experience with children, and attitudes toward children would be related to adults’ judgments. Although men had significantly higher confidence-of-event-occurrence scores than women, significant gender differences in  $d'$  and  $\beta$  failed to emerge, indicating comparable sensitivity and bias. Instead, age (being older) predicted greater accuracy and a more conservative response bias when children denied that the event occurred. This was not explained by a difference in experience with children, attitudes toward child witnesses, or participant type (student vs. layperson). Rather, with age and maturity, adults may better appreciate (and refine their judgments of) children’s denials. The other individual differences were not predictive of ability to distinguish children’s reports.

### Caveats

Several caveats accompany our conclusions. First, although some of the stimuli events (such as being punished or embarrassed) were negative events, few were prosecutable. Children’s demeanor and reporting styles may differ for prosecutable events, such as abuse, perhaps especially if the children believe that someone could face serious negative consequences as a result of their report. Second, the leading and repetitious nature of our interviews may have influenced our findings. For example, regarding false denials, the fact that children denied the events even in the face of leading questions may have implied to observers that the children must have been accurately denying the event’s occurrence. Third, we relied on parental report for “ground truth,” but parents’ memories can be fallible. Fourth, observers were informed that the children had been instructed not to lie. This feature bolstered generalizability to forensic interviews and courtroom settings where children also would be instructed to tell the truth and warned against lies but not to situations where observers are uninformed about this instruction. Fifth, future research should include a larger number of laypersons to test for possible participant type differences in discernment, and additional research should examine how *groups* make decisions of the sort we studied. Judgment errors made by individuals acting alone could be corrected during the group deliberation process that occurs in jury trials, parole board hearings, and other collective decision-making contexts. Sixth, although several random orders of stimuli were used, fatigue could have influenced our results. However, fatigue effects would not exceed what one would expect in various legal contexts, such as a trial. Moreover, although ratings were obtained on a relatively large number of children, the number was too small to examine several potentially important child factors (e.g., age, gender). Further research on such factors is needed.

Seventh, the Child Witness Questionnaire was always presented after the videotaped interviews. The interviews may have affected responses on this measure. Eighth, in the legal context, fact finders may at times be aware of reasons for children to falsely deny (e.g., fear of negative consequences for the children and/or the perpetrator) and thus evince appropriate skepticism of such responses. Finally, additional research should include children from a wider range of ages.

### Conclusion

Overall, our findings suggest that, when evaluating leading interviews of preschool children, adults have particular difficulty identifying false denials. The problem of false denials has received relatively little research attention (but see Lyon, 1999; Pipe & Goodman, 1991) compared with the problem of false reports (Ceci & Bruck, 1993), the latter of which is often discussed by social scientists and legal scholars alike. In the criminal law context, this focus may be appropriate given the assumption of the United States legal system that it is worse to imprison an innocent person than to set a guilty person free. However, the failure to bring criminals to justice is also a danger to society, especially when those criminals are likely to reoffend and when the societal and human costs of the offenses are high. Furthermore, the failure to identify false denials can affect not only criminal cases but other types of cases as well, including custody violations (“No, daddy

didn't take me out of state”), breach of contract (“No, the repairman didn't use the tools that you said he shouldn't use”), or medical malpractice (“No, the nurse didn't ask me if I had eaten breakfast before giving me the needle”). These negative consequences could hopefully be mitigated by increased awareness and policy changes aimed at improving fact finders' abilities to evaluate all types of reports.

References

Ackil, J., & Zaragoza, M. (1998). Memorial consequences of forced confabulation: Age differences in susceptibility to false memories. *Developmental Psychology, 34*, 1358–1372. doi:10.1037/0012-1649.34.6.1358

Ball, C. T., & O'Callghan, J. (2001). Judging the accuracy of children's recall: A statement-level analysis. *Journal of Experimental Psychology: Applied, 7*, 331–345. doi:10.1037/1076-898X.7.4.331

Bjorklund, D., Cassel, W., Bjorklund, B., Brown, R., Park, C., Ernst, K., . . . Owen, F. A. (2000). Social demand characteristics in children's and adults' eyewitness memory and suggestibility: The effect of different interviewers on free recall and recognition. *Applied Cognitive Psychology, 14*, 421–433. doi:10.1002/1099-0720(200009)14:5<421::AID-ACP659>3.0.CO;2-4

Bottoms, B. L., Golding, J. G., Stevenson, M., Wiley, T., & Wozniak, J. (2006). A review of factors affecting jurors' decisions in child sexual abuse cases. In M. Toglia, J. D. Read, D. F. Ross, & R. C. L. Lindsay (Eds.), *The psychology of eyewitness memory: Vol. 1: Memory for events* (pp. 509–544). Mahwah, NJ: Erlbaum.

Brigham, J. C. (1998). Adults' evaluations of characteristics of children's memory. *Journal of Applied Developmental Psychology, 19*, 15–39. doi:10.1016/S0193-3973(99)80026-4

Bruck, M., Ceci, S., & Hembrooke, H. (1998). Reliability and credibility of young children's reports: From research to policy and practice. *American Psychologist, 53*, 136–151. doi:10.1037/0003-066X.53.2.136

Bussey, K. (1992). Lying and truthfulness: Children's definitions, standards, and evaluative reactions. *Child Development, 63*, 129–137. doi:10.2307/1130907

Cassel, W., Roebbers, C., & Bjorklund, D. (1996). Developmental patterns of eyewitness responses to repeated and increasingly suggestive questions. *Journal of Experimental Child Psychology, 61*, 116–133. doi:10.1006/jecp.1996.0008

Castelli, P., Goodman, G. S., & Ghetti, S. (2005). Effects of interview style and witness age on perceptions of children's credibility in sexual abuse cases. *Journal of Applied Social Psychology, 35*, 297–319. doi:10.1111/j.1559-1816.2005.tb02122.x

Ceci, S. J., & Bruck, M. (1993). Suggestibility of the child witness: A historical review and synthesis. *Psychological Bulletin, 113*, 403–439. doi:10.1037/0033-2909.113.3.403

Ceci, S. J., & Huffman, M. (1997). How suggestible are preschool children? *Journal of the American Academy of Child & Adolescent Psychiatry, 36*, 948–958. doi:10.1097/00004583-199707000-00017

Ceci, S. J., Loftus, E. F., Leichtman, M., & Bruck, M. (1994). The possible role of source misattributions in the creation of false beliefs among preschoolers. *International Journal of Clinical and Experimental Hypnosis, 42*, 304–320. doi:10.1080/00207149408409361

Chahal, K., & Cassidy, T. (1995). Deception and its detection in children: A study of adult accuracy. *Psychology, Crime & Law, 1*, 237–245. doi:10.1080/10683169508411959

Diamond, S., & Dolan, H. J. (2010). Achieving diversity on the jury. *Researching Law, 21*, 1–6.

Edelstein, R. S., Luten, T. L., Ekman, P., & Goodman, G. S. (2006). Detecting lies in children and adults. *Law and Human Behavior, 30*, 1–10. doi:10.1007/s10979-006-9031-2

Fivush, R. (1993). Developmental perspectives on autobiographical recall.

In G. S. Goodman & B. L. Bottoms (Eds.), *Child victims, child witnesses* (pp. 1–24). New York, NY: Guilford Press.

Ghetti, S., & Alexander, K. (2004). “If it happened, I would remember it”: Strategic use of event memorability in the rejection of false autobiographical events. *Child Development, 75*, 542–561. doi:10.1111/j.1467-8624.2004.00692.x

Goodman, G. S. (1984). The child witness: Conclusions and future directions for research and legal practice. *Journal of Social Issues, 40*, 157–175. doi:10.1111/j.1540-4560

Goodman, G. S., Myers, J., Qin, J., Quas, J. A., Castelli, P., Redlich, A., & Rogers, L. (2006). Hearsay versus children's testimony: Effects of truthful and deceptive statements on jurors' decisions. *Law and Human Behavior, 30*, 363–401. doi:10.1007/s10979-006-9009-0

Goodman, G. S., & Reed, R. S. (1986). Age differences in eyewitness testimony. *Law and Human Behavior, 10*, 317–332. doi:10.1007/BF01047344

Goodman-Brown, T. B., Edelstein, R. S., Goodman, G. S., Jones, D. P. H., & Gordon, D. S. (2003). Why children tell: A model of children's disclosure of sexual abuse. *Child Abuse & Neglect, 27*, 525–540. doi:10.1016/S0145-2134(03)00037-1

Green, D., & Swets, J. (1966). *Signal detection theory and psychophysics*. New York, NY: Wiley.

Harris, L., Goodman, G. S., Augusti, E.-M., Chae, Y., & Alley, D. (2009). Children's resistance to suggestion. In K. Kuenhle & M. Connell (Eds.), *The evaluation of child sexual abuse allegations* (pp. 181–202). New York, NY: Wiley.

Heaps, C. M., & Nash, M. (2001). Comparing recollective experience in true and false autobiographical memories. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27*, 920–930. doi:10.1037/0278-7393.27.4.920

Hyman, I. E., Jr., & Pentland, J. (1996). The role of mental imagery in the creation of false childhood memories. *Journal of Memory and Language, 35*, 101–117. doi:10.1006/jmla.1996.0006

Laimon, R. L., & Poole, D. A. (2008). Adults usually believe young children: The influence of eliciting questions and suggestibility presentations on perceptions of children's disclosures. *Law and Human Behavior, 32*, 489–501. doi:10.1007/s10979-008-9127-y

Landström, S., Granhag, P. A., & Hartwig, M. (2007). Children's live and videotaped testimonies: How presentation mode affects observers' perception, assessment, and memory. *Legal and Criminal Psychology, 12*, 333–347. doi:10.1348/135532506X133607

Leach, A., Talwar, V., Lee, K., Bala, N., & Lindsay, R. (2004). “Intuitive” lie detection of children's deception by law enforcement officials and university students. *Law and Human Behavior, 28*, 661–685. doi:10.1007/s10979-004-0793-0

Lee, K., & Cameron, C. (2000). Extracting truthful information from lies: Emergence of the expression–representation distinction. *Merrill-Palmer Quarterly, 46*, 1–20.

Leichtman, M. D., & Ceci, S. J. (1995a). The effects of stereotypes and suggestions on preschoolers' reports. *Developmental Psychology, 31*, 568–578. doi:10.1037/0012-1649.31.4.568

Leichtman, M. D., & Ceci, S. J. (1995b). The effects of stereotypes and suggestions on preschoolers' reports: Correction. *Developmental Psychology, 31*, 758. doi:10.1037/0012-1649.31.5.758

Leippe, M. R., Manion, A. P., & Romanczyk, A. (1993). Discernability or discrimination? Understanding jurors' reactions to accurate and inaccurate child and adult eyewitnesses. In G. S. Goodman & B. L. Bottoms (Eds.), *Child victims, child witnesses* (pp. 169–201). New York, NY: Guilford Press.

Loftus, E. F., & Pickrell, J. E. (1995). The formation of false memories. *Psychiatric Annals, 25*, 720–725.

Lyon, T. D. (1999). The new wave in children's suggestibility research: A critique. *Cornell Law Review, 84*, 1004–1087.

Lyon, T. D., Malloy, L., Quas, J., & Talwar, V. (2008). Coaching, truth

- induction, and young maltreated children's false allegations and false denials. *Child Development*, 79, 914–929. doi:10.1111/j.1467-8624.2008.01167.x
- MacMillan, N., & Creelman, C. D. (1991). *Detection theory*. New York, NY: Cambridge University Press.
- Magnussen, S., Andersson, S., Cornoldi, C., De Beni, R., Endestad, T., Goodman, G. S., . . . Zimmer, H. (2006). What people believe about memory. *Memory*, 14, 595–613. doi:10.1080/09658210600646716
- Orcutt, H. K., Goodman, G. S., Tobey, A. E., Batterman-Faunce, J. M., & Thomas, S. (2001). Detecting deception in children's testimony: Fact-finders' abilities to reach the truth in open court and closed-circuit trials. *Law and Human Behavior*, 25, 339–372. doi:10.1023/A:1010603618330
- Ost, J., Vrij, A., Costall, A., & Bull, R. (2002). Crashing memories and reality monitoring: Distinguishing between perceptions, imaginings and false memories. *Applied Cognitive Psychology*, 16, 125–134. doi:10.1002/acp.779
- Pipe, M.-E., & Goodman, G. S. (1991). Elements of secrecy: Implications for children's testimony. *Behavioral Sciences and the Law*, 9, 33–41. doi:10.1002/bsl.2370090105
- Polak, A., & Harris, P. (1999). Deception by young children following noncompliance. *Developmental Psychology*, 35, 561–568. doi:10.1037/0012-1649.35.2.561
- Poole, D. A., & Lindsay, D. S. (1998). Assessing the accuracy of young children's reports: Lessons from the investigation of child sexual abuse. *Applied and Preventive Psychology*, 7, 1–26. doi:10.1016/S0962-1849(98)80019-X
- Qin, J., Ogle, C. M., & Goodman, G. S. (2008). Adults' memories of childhood: True and false reports. *Journal of Experimental Psychology: Applied*, 14, 373–391. doi:10.1037/a0014309
- Reyna, V., Holliday, R., & Marche, T. (2002). Explaining the development of false memories. *Developmental Review*, 22, 436–489. doi:10.1016/S0273-2297(02)00003-5
- Schaaf, J. M., Alexander, K. W., & Goodman, G. S. (2008). Children's false memory and true disclosure in the face of repeated questions. *Journal of Experimental Child Psychology*, 100, 157–185. doi:10.1016/j.jecp.2007.09.002
- Schooler, J. W., Gerhard, D., & Loftus, E. F. (1986). Qualities of the unreal. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 12, 171–181. doi:10.1037/0278-7393.12.2.171
- Strömwall, L., & Granhag, P. (2005). Children's repeated lies and truths: Effects on adults' judgments and reality monitoring scores. *Psychiatry, Psychology, and Law*, 12, 345–356. doi:10.1375/pplt.12.2.345
- Strömwall, L. A., Granhag, P., & Landström, S. (2007). Children's prepared and unprepared lies: Can adults see through their strategies? *Applied Cognitive Psychology*, 21, 457–471. doi:10.1002/acp.1288
- Talwar, V., Gordon, H., & Lee, K. (2007). Lying in the elementary school years: Verbal deception and its relation to second-order belief understanding. *Developmental Psychology*, 43, 804–810. doi:10.1037/0012-1649.43.3.804
- Talwar, V., & Lee, K. (2002). Development of lying to conceal a transgression: Children's control of expressive behavior during verbal deception. *International Journal of Behavioral Development*, 26, 436–444. doi:10.1080/01650250143000373
- Talwar, V., & Lee, K. (2008). Social and cognitive correlates of children's lying behavior. *Child Development*, 79, 866–881. doi:10.1111/j.1467-8624.2008.01164.x
- Talwar, V., Murphy, S., & Lee, K. (2007). White lie-telling in children for politeness purposes. *International Journal of Behavioral Development*, 31, 1–11. doi:10.1177/0165025406073530
- Vrij, A., Edward, K., & Bull, R. (2001). Stereotypical verbal and nonverbal responses while deceiving others. *Personality and Social Psychology Bulletin*, 27, 899–909. doi:10.1177/0146167201277012
- Wells, G. L., & Windschitl, P. D. (1999). Stimulus sampling and social psychological research. *Personality and Social Psychology Bulletin*, 25, 1115–1125. doi:10.1177/01461672992512005
- Wrightsmann, L. S., Batson, A. L., & Edkins, V. A. (Eds.). (2004). *Measures of legal attitudes*. Belmont, CA: Wadsworth.

## Appendix

### Sample List of Interview Questions

The first and last questions were always asked in the first and last position and in the exact wording below. The order of the other questions was randomly varied across interviews, and the wording was customized to match the event about which the child was questioned (e.g., another kid pulled down your pants in front of a store).

What happened the time [another kid pulled down your pants in front of a store]?

Was your mom there when [the other kid pulled down your pants in front of the store]?

Did you cry when [the kid pulled down your pants by the front of the store]?

What did you do that time [another kid pulled down your pants in front of a store]?

Who was with you the time [another kid pulled down your pants in front of the store]?

What else can you remember about [the time another kid pulled down your pants in front of a store]?

Received October 9, 2009

Revision received August 2, 2010

Accepted June 29, 2011 ■