

# The Effects of Stereotypes and Suggestions on Preschoolers' Reports

Michelle D. Leichtman  
Harvard University

Stephen J. Ceci  
Cornell University

Children's ( $N = 176$ ) reported memories of a strange man's visit were studied. Three- to 6-year-olds were interviewed repeatedly after the event in one of the following conditions: (a) *control*, in which no interviews contained suggestive questions; (b) *stereotype*, in which children were given previsit expectations about the stranger; (c) *suggestion*, in which interviews contained erroneous suggestions about misdeeds committed by the stranger; and (d) *stereotype plus suggestion*, in which children were given both pre- and postvisit manipulations. Results from open-ended interviews after 10 weeks indicated that control participants provided accurate reports, stereotypes resulted in a modest number of false reports, and suggestions resulted in a substantial number of false reports. Children in the stereotype-plus-suggestion group made high levels of false reports. All experimental conditions showed dramatic developmental trends favoring older children.

A burgeoning literature on children's suggestibility has appeared over the past decade, spawned by both theoretical issues surrounding memory development and applied issues surrounding children's courtroom testimony. Put simply, the theoretical issues in children's suggestibility concern one or more of a family of cognitive and social developmental factors, whereas the applied issues concern the limits of children's testimonial competence and, particularly, the issue of ecological validity. In the present study, we build on a base of work from both arenas to address the issue of how two critical factors might affect preschool children's reports of an event that centers on the actions of a particular person. These factors are the stereotypes about a person held by children before their witnessing the event of interest and repeated suggestive questioning that occurs during multiple interviews after the event.

To put our work in context, we briefly indicate below the cognitive and social factors that have received the most attention from researchers concerned with the suggestibility of children's memory and reporting. We then discuss the related testimonial issues.

## Cognitive Issues

In a recent review of research on children's suggestibility, 83% of studies comparing preschool-age children's suggestibility with older children's are reported to have found larger suggestibility effects for the former (see Table 2 in Ceci & Bruck, 1993). Although the underlying mechanisms responsible for these age-related differences have yet to be determined, several cognitive and social developmental candidates have been

touted. The most actively researched candidate has been developmental differences in trace strength that may make young children's memories especially vulnerable to featural disintegration or resistant to retrieval-time relearning (e.g., Brainerd & Reyna, 1988; Brainerd, Reyna, Howe, & Kingma, 1990; Howe, 1991). As conceptualized by some theorists, the incorporation of postevent information occurs as a function of the strength of the memory trace, with weak traces being especially vulnerable to featural dilution or blending (i.e., *destructive updating*) or total dissolution or erasure (Ceci, Toglia, & Ross, 1988). One important prediction of trace theory is that age differences in suggestibility will occur because younger children encode weaker (and less "gistlike") traces, which are more vulnerable to featural disintegration or overwriting compared with those of their older counterparts. Although a number of studies lend support to this position (King & Yuille, 1987; Warren, Hulse-Trotter, & Tubbs, 1991), some theorists have also challenged the notion that suggestibility is related to trace strength (Howe, 1991; Zaragoza, Dahlgren, & Muench, 1992).

Another candidate that has been put forward as a possible cause of developmental differences in suggestibility is *source misattributions*. Source misattributions are found when participants are asked either to perform some act or else to merely imagine performing it. Often, younger children are more likely to misattribute performing an act that they had only imagined performing (e.g., Foley & Johnson, 1985; Foley, Johnson, & Raye, 1983; Lindsay, Johnson, & Kwon, 1991). Although there are no studies of the role of source misattributions in suggestibility, it seems plausible that they may play a role. It may be that the effect of repeated suggestive questions is to induce the participant to create images of the suggested events; later, participants may have difficulty distinguishing the sources of their stored images—are they based on direct perception or merely on internally generated imagery? On the basis of the developmental findings just noted, it could be that very young children's suggestibility is a result of their source confusions.

## Social Factors

Bribes, threats, fear of embarrassment, protection of loved ones, and desire to gain material rewards all have been shown

---

Michelle D. Leichtman, Department of Psychology, Harvard University; Stephen J. Ceci, Department of Human Development, Cornell University.

This research was supported by Grant RO1 HD 25775 from the National Institute of Child Health and Human Development.

Correspondence concerning this article should be addressed to Michelle D. Leichtman, Department of Psychology, Harvard University, 33 Kirkland Street, Cambridge, Massachusetts 02138.

to influence preschoolers' report accuracy (e.g., Bussey, 1992; Ceci, Leichtman, & Putnick, 1992; Lewis, Stranger, & Sullivan, 1989; Peters, 1991). A precise comparison between the level of distortion that children display in their reports in the presence of these external pressures and the level of distortion of adults under similar conditions has not been made in the literature. However, broadly speaking, such motivations appear to affect both children and adults, and a differential sensitivity to general demand characteristics suggests that young children may be comparatively more prone to report distortion in the face of these motivations under some circumstances. For example, preschoolers have been shown to be more suggestible than older children, particularly when erroneous suggestions are made by adult authority figures (Ceci, Ross, & Tolia, 1987), which may be a result of their greater tendency to consider their own memories less reliable than an adult's or their desire to please adult interviewers by telling them what the child imagines they want to hear. (Additional social-motivational factors are discussed later, in connection with testimonial issues.)

To recap, observed developmental differences in suggestibility may be the result of cognitive factors such as age differences in trace strength and source misattributions, and social factors such as bribes, threats, and expectations. In addition to the basic research conducted on these issues, there are many studies that have been animated by a desire to learn more about how these factors conspire to influence children's statements to forensic interviewers and their testimony in court (e.g., Goodman et al., 1992). We turn to this next, as a means of framing the present experiment.

### Testimonial Issues

Despite the widespread empirical evidence of age differences in suggestibility, it remains unclear whether similar levels of suggestibility may be assumed to be present in real-world cases involving the testimony of individuals in the courtroom. The overarching concern here is that many experiments, in particular those from the first half of this century (see Ceci & Bruck, 1993), lack sufficient ecological validity to allow us to confidently extrapolate their findings to the real world. Specifically, the experimental conditions in which participants demonstrate their memorial suggestibility under the watchful eyes of researchers differ in several significant ways from those in which individuals give forensically relevant testimony. The affectively laden nature of the encoding and retrieval contexts in some actual forensic situations, as well as the motivational forces and demand characteristics involved in these situations, is difficult to ethically incorporate into empirical research programs (Ceci, Leichtman, & Bruck, in press). As a result, we still do not know much about the way children respond to them.

The goal of the present study was to evaluate suggestibility under conditions that have not heretofore been investigated, but that nevertheless characterize a large number of situations in which child witnesses eventually appear in court. Because these conditions have gone unexplored, it is not clear whether the suggestibility effects documented in the existing literature underestimate or overestimate the magnitude of children's reported distortion under these real-world forensic conditions. Certain factors that frequently crop up in courtroom testimony,

particularly those pertaining to the number and timing of witness interviews, would lead one to suspect that youngsters relating past events are more prone to suggestion in such situations than current research would have us believe. Alternatively, other variables that play a role in the real world, such as the salience of events about which children testify, would lead to the opposite conclusion, namely, that experimental conditions give rise to less optimistic views of children's resistance to suggestion than those reflected by their real-world analogs. In the present work, we considered how a number of contextual factors involved in the experience and reporting of an event might affect the accuracy of preschoolers' reports. Following this, we evaluated the ability of condition-blind adults to assess whether children's reports were factually accurate. We did so because adults' ability to determine the accuracy of children's reports bears heavily on the importance of the issue of suggestibility in legal cases.

### Stereotypes as Person Schemas

Stereotypes are naive theories about personal characteristics, which function to organize and structure experience by directing individuals to look for expectancies in their environment and advising them on how to interpret such expectancies. Thus, stereotypes are a form of schematic knowledge that help organize memory, by adding thematically congruent information that was not perceived, or sometimes by distorting what is perceived (Martin & Halverson, 1983; Strangor & McMillan, 1992).

Before witnessing an event, a child may be provided with a particular stereotype about the person involved, and this may direct the child's attention to expectancy-congruent behaviors (e.g., in court cases, the defendant might be an estranged parent who has been previously criticized by the custodial parent in the child's presence, and the child may even have come to accept these criticisms as stable aspects of the parent's character). Hence, such behaviors may be remembered and reported disproportionately.

### Repeated Suggestions Over Long Intervals

It is probably not an exaggeration to say that the presence of multiple repetitive interviews over the course of long retention intervals has become the norm in cases in which children testify (see Ceci & Bruck, 1993, for similar examples from other cases; see Humphrey, 1985).<sup>1</sup> Recent estimates indicate that by the time they get to court, children have been subjected to 4 to 11 forensic interviews, on average, and in most cases they have ex-

<sup>1</sup> In 1985, the attorney general for Minnesota criticized the repeated interviewing of the child witnesses in the so-called "Jordan case" (Humphrey, 1985), which, along with the McMartin case (*People v. Buckley*, 1984), was the first and the best known of the mass allegation child sexual abuse cases: "[O]ne child had already been interviewed by nine individuals about the alleged abuse. The mother of another child indicated that her daughter had been interviewed at least thirty and possibly as many as fifty times by law enforcement or Scott County authorities. A large number of other children also were repeatedly interviewed" (Humphrey, 1985, p. 9).

perienced numerous other bouts of questioning from family members, therapists, social workers, and other interested parties (Gray, 1993; McGough, 1994). Nonetheless, most studies to date have focused on the suggestibility of children after a single suggestive interview, as reflected in Ceci and Bruck's (1993) comprehensive review. (For exceptions, see, for example, Goodman & Clarke-Stewart, 1991; Lepore & SESCO, 1994; Poole & White, 1991, 1993.)

#### Texas v. Macias (1987)

A 1987 death row case in El Paso, Texas, illustrates how the confluence of testimonial factors we have referred to could influence the outcome of a case. The most important testimony at the original trial in this case was provided by a child who claimed that on one particular occasion, she remembered seeing the defendant with blood on his shirt and a weapon in his hand at his trailer. The child's mother had told her on numerous occasions that the defendant was a bad man and had warned her against being friendly with him, providing a schema relating to his character before the event she had allegedly witnessed. From interviews with the child, it was clear that she possessed a deeply held negative stereotype about the defendant, and her statements to the police as well as her trial testimony were in concert with this stereotype. In addition, the child witness in the Macias case was subjected to relentless inquiries in pursuit of her event-relevant memories, in a series of highly suggestive interviews that stretched over an extended period of time.

This legal case offers a window through which to consider one possible outcome of a situation in which a child is provided with both a preexisting negative stereotype about a particular defendant and a large number of suggestive interviews in line with this stereotype. In this case, several years after incriminating the defendant in her testimony, the key child witness gave a sworn deposition recanting her testimony. In this statement, she said that the repeated interviews had confused her and that she had said things that were incorrect, because she had wanted to help the adults involved and had known that the defendant was a bad man. In her words:

When I first saw Fred with red stuff on his shirt, I didn't think it was that important. At first, I didn't really know if it was blood or chili. [Note: the defendant worked in a salsa factory, as did Jennifer's mother, hence, the sight of red on people's clothing was common.] Later, when I saw Fred had a gun or a knife, it caught my attention, and I thought that it must be blood. Because different people asked me so many different questions about what I saw, I became confused. I thought I might have seen something that would be helpful to the police. I didn't realize that it would become so important. I thought they wanted me to be certain, so I said I was certain even though I wasn't. Originally, I think I told the police just what I saw. But the more questions I was asked, the more confused I became. I answered questions I wasn't certain about because I wanted to help the adults. (*Texas v. Macias*, 1987; subscribed and sworn before Regina Jarius, Notary, on the 13th day of August, 1988)

This deposition, made 12 days before the defendant's scheduled execution, resulted in a stay of execution, and the defendant was ultimately freed.

The present study was designed to experimentally examine

the combined effect of stereotypes and repeated suggestive interviews. We refer to it as the "Sam Stone Study," because the event of interest was the visit of a man named Sam Stone to the day-care centers of our participants.

## Method

### Participants

One hundred and seventy-six preschoolers participated in this experiment. They were enrolled in private day-care centers, and they represented a wide range of social and ethnic groups, with approximately 15% of all participating children coming from families receiving Aid to Families With Dependent Children (AFDC) and the remaining children coming about evenly split between blue-collar/middle-class and white-collar/professional families. The children were divided into two age groups: early preschoolers (3- and 4-year-olds) and older preschoolers (5- and 6-year-olds). Assignment to experimental condition was random, but the unit of assignment was the classroom ( $n = 8$ ) rather than the individual child. (This was done to obviate the potential contaminating effect of classmates sharing with each other the details of their interviews.)

### Procedure

Children were assigned to one of four conditions, denoted as follows: (a) control, (b) stereotype, (c) suggestion, and (d) stereotype plus suggestion. The central event of interest was the visit of a stranger named Sam Stone to the preschoolers in all conditions at their day-care centers. In each of the eight day-care classrooms, Sam Stone enacted the same scripted event. First, he entered the classroom and said hello to a teacher or aid who sat amidst the assembled children during a story-telling session, and he was introduced by the teacher or aid to the children. Next, he commented on the story that was being read to the children by the teacher or aid ("I know that story; it's one of my favorites!") and strolled around the perimeter of the classroom. Finally, he departed, waving goodbye to the children. In each case, the entire event was timed and lasted approximately 2 min.

Two experimental manipulations, a preevent and a postevent manipulation, formed the basis of the differences among the four conditions. All of the children, including those in the three experimental groups as well as the controls, received a forensic interview approximately 10 weeks after Sam Stone's visit. However, children in the *control* group received no information about Sam Stone before his visit and were questioned once a week during the 4 weeks immediately following this visit in a neutral manner. That is, during the four interviews, control children were simply asked questions about what Sam Stone had done during his visit to their school and were given no suggestions about the nature of this visit or Sam Stone's activities.

Children in the *stereotype* condition, in contrast, received considerable information about Sam Stone's personality before his visit to their school. Each week, beginning a month before the visit, research assistants went to the children's day-care centers, and in the course of playing with them, presented 3 different scripted stories about Sam Stone (for a total of 12 stories over the four visits; see Appendix A). In each of these stories, Sam Stone was depicted as a kind, well-meaning, but very clumsy and bumbling person. For example:

You'll never guess who visited me last night. [pause] That's right. Sam Stone! And guess what he did this time? He asked to borrow my Barbie and when he was carrying her down the stairs, he accidentally tripped and fell and broke her arm. That Sam Stone is always getting into accidents and breaking things! But it's okay, because Sam Stone is very nice and he is getting my Barbie doll fixed for me.

Following Sam Stone's visit, children in the stereotype condition were treated identically to the control group, receiving four neutral interviews over the 4 weeks following his visit, and a fifth interview 10 weeks after the visit.

A third group of children, those in the *suggestion* condition, did not receive the preevent manipulation just described (i.e., the stereotype induction) but did receive a postevent manipulation consisting of suggestive interviews following their encounter with Sam Stone. Thus, although this group of children had no knowledge of Sam Stone before his visit to their classes, they received four interviews following his visit that were quite different from the neutral interviews given to children in the control and stereotype conditions. During their interviews, children in the suggestion group were provided with two erroneous suggestions about what occurred during Sam Stone's visit, embedded within an interview that was otherwise parallel to those of the control and stereotype groups. The first misleading suggestion was that Sam Stone had ripped a book, and the second was that he had soiled a teddy bear. The exact questions about the events that occurred during Sam Stone's visit were different for this group of children during each of the interviews, but the same implications were embedded in each. For example, 1 week children were asked, "When Sam Stone got that bear dirty, did he do it on purpose or was it an accident?," and in the following interview session they were asked, "Was Sam Stone happy or sad that he got that bear dirty?" (See Appendix B for details of interview protocols and questions.)

A fourth group of children, those in the *stereotype-plus-suggestion* condition, were exposed to both the preevent stereotype and the postevent leading question manipulations. These children were thus provided with misleading information about Sam Stone at two points in the process of acquiring information about him that could bear on their later reports.

The fifth interview, experienced by all children, was conducted by a new interviewer, who was not present during Sam Stone's visit or the first four interviews. In this case, the same questions were asked and the same forensic procedures were used to interrogate children in all groups (see Appendix C). In each case, children were first made to feel comfortable; a free narrative was then elicited from them ("Remember the day that Sam Stone visited your school? Well, I wasn't there that day, and I'd like you to tell me everything that happened when he visited"); and, finally, they were given probing questions about specific events. These specific probes were directed at the two events that did not occur during Sam Stone's visit but that children in the suggestion and stereotype-plus-suggestion conditions had heard about before, namely, Sam Stone's soiling a teddy bear and ripping a book. These probe questions asked children whether they had "heard something" about the items and whether they had seen Sam Stone engage in some activity with them. Thus, our central analyses focus on children's responses to the free narrative, as well as their initial responses to these probes.

In addition, for only those children whose answers to the probes indicated that they actually saw Sam Stone commit nonevents, countersuggestion questions were posed, to attempt to gauge the strength of their statements (e.g., "You didn't really see him do this, did you?"). In the following section, we report for each of our experimental conditions first the free narrative and probe data, and then results for the subset of children asked the countersuggestion questions.

## Results and Discussion

Data from the fifth interview were coded from videotape by condition-blind raters, who categorized children's answers in terms of their content, scoring "don't know" and "no response," as well as specific details provided by children. Twenty percent of the videotaped interviews were randomly selected and recoded by an independent rater, and interrater reliability

### CONTROL GROUP (NO STEREOTYPE; NO SUGGESTIONS)

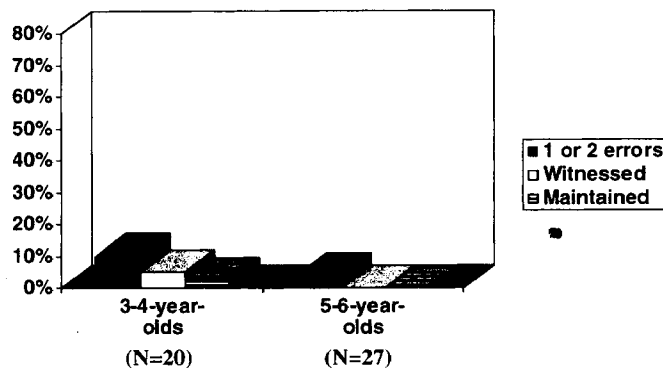


Figure 1. Percentage of preschoolers' answers that were erroneous. Light-colored bar indicates that the child asserted that an incorrect event occurred; dark-colored bar indicates that the child claimed to have actually observed the nonevent; hash-marked bar indicates that the child insisted on having witnessed the event, despite mild attempt at dissuading.

was found to be high (Cohen's  $\kappa = .90$ ). Below, we begin by describing children's accuracy in each of the four experimental conditions, and then we proceed to subject these data to various statistical analyses as a function of age and condition. For the purposes of the first set of data analyses, we consider the raw (unconditionalized) percentages of children who both reported that Sam Stone engaged in the nonevent and resisted the countersuggestion.

### Control Condition

No child in the control group made any false allegations in his or her free narratives when initially asked by the interviewer during the fifth interview to tell everything they could remember about the day that Sam Stone visited their classroom. As seen in Figure 1, nearly all of the 47 children assigned to the control group resisted claiming anything erroneous had occurred not only in their free narratives but also in response to probes. Thus, when specifically probed about a book or teddy bear, only 10% of the youngest control group children's claims indicated that Sam Stone did anything to a book or teddy bear (i.e., 4 claims out of 40 opportunities). Furthermore, when specifically asked if they actually saw him do anything to a book or teddy bear, as opposed to merely hearing that he did something, only 5% of the younger preschoolers' claims continued to indicate that anything occurred (i.e., 2 claims out of 40 opportunities). Finally, when gently challenged with countersuggestions such as "You didn't really see him do anything to the book (the teddy bear), did you?," only 2.5% of the younger children's claims (1 out of 40 opportunities) indicated that they actually observed him doing so. In summary, in the absence of any attempt by adults to taint the youngest children's reports before the fifth and final interview, their reports were largely, although not wholly, void of errors. And yet these children's reports usu-

### STEREOTYPE GROUP (STEREOTYPE; NO SUGGESTIONS)

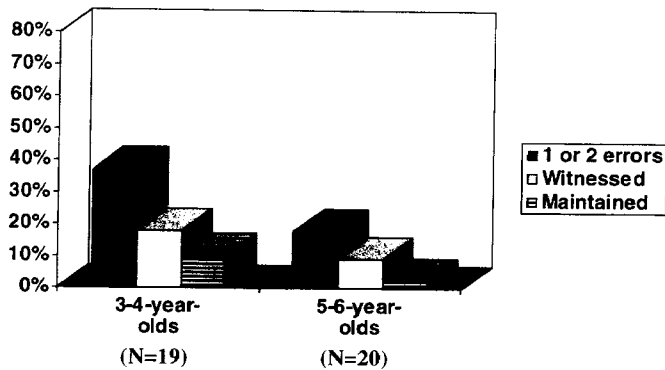


Figure 2. Percentage of preschoolers' answers that were erroneous. Light-colored bar indicates that the child asserted that an incorrect event occurred; dark-colored bar indicates that the child claimed to have actually observed the nonevent; hash-marked bar indicates that the child insisted on having witnessed the event, despite mild attempt at dissuading.

ally included accurate accounts of actual information; they often were able to recall Sam Stone's limited activities on the day he visited, for example, that he walked around the housekeeping section of the classroom, that he greeted the children pleasantly, or that he waved goodbye. As for the older children, no child made any false allegations in his or her free narratives, and only 2 of their claims in response to initial probes (out of 54 possible) indicated that Sam Stone committed a misdeed, and both of these were readily relinquished when the children were asked if they had actually witnessed the misdeed themselves. Because none of the older children claimed to have actually observed Sam Stone damage either item, they were not asked countersuggestions to see if they would relinquish their erroneous claims. Like their younger counterparts, these older children's recall was filled with examples of actual events that actually occurred.

#### Stereotype Condition

As was the case in the control condition, none of the children assigned to the stereotype condition claimed that they observed Sam Stone damaging either item in their free narrative, when they were initially asked by the interviewer during the fifth interview to tell everything they could remember about Sam Stone's visit. As seen in Figure 2, however, the stereotyping manipulation did have an effect on probed recall, particularly for the youngest children. In the final interview with this group, in response to the probes, "Did Sam Stone rip the book (soil the teddy bear)?," 37% of their responses indicated that he did at least one of these things. Of these children, 18% subsequently claimed they saw Sam Stone do these misdeeds (7 out of 38 opportunities). But, after being gently challenged, only 10% of their responses continued to indicate that they witnessed him do these things. In contrast to younger children, older preschoolers were significantly more resistant to the influence of

the stereotype, with roughly half the rate of errors at all three levels of probing. Only 1 older child, in a single response (out of 40 opportunities), continued to indicate he had seen Sam Stone commit a misdeed after being gently challenged with a counter-suggestion ("He didn't really do this, did he?").

#### Suggestion Condition

Unlike the control and stereotype conditions, some children assigned to the suggestion condition claimed that they observed Sam Stone damaging either item in their free narrative. Twenty-one percent of the youngest children (6 out of 29) and 14% of the older children (3 out of 22) made spontaneous claims in their free narratives regarding damaged books (or teddy bears or both). This finding is rare in the literature, as children's suggestibility is usually confined to cued recall and recognition measures, with few if any errors in free recall (Ceci & Bruck, 1993).

As can be seen in Figure 3, in response to the probe questions, 53% of the youngest children's responses in the suggestion condition and 38% of the older children's indicated that Sam Stone did one or both misdeeds. Moreover, in response to follow-up probes, 35% of the youngest children's responses indicated that they had actually seen him do these things, as opposed to being told he did them. Finally, even after being challenged with the countersuggestion, 12% of the youngest children continued to claim they saw him do one or both misdeeds. Older children were also susceptible to the suggestive interviews, though at very reduced levels: Ultimately, only 2 out of 22 of the older children continued to maintain that they saw him do the misdeeds when challenged with a countersuggestion.

#### Stereotype-Plus-Suggestion Condition

Finally, in the stereotype-plus-suggestion condition, 46% of the youngest children and 30% of the oldest children spontane-

### STEREOTYPE PLUS SUGGESTION GROUP (STEREOTYPE; SUGGESTIONS)

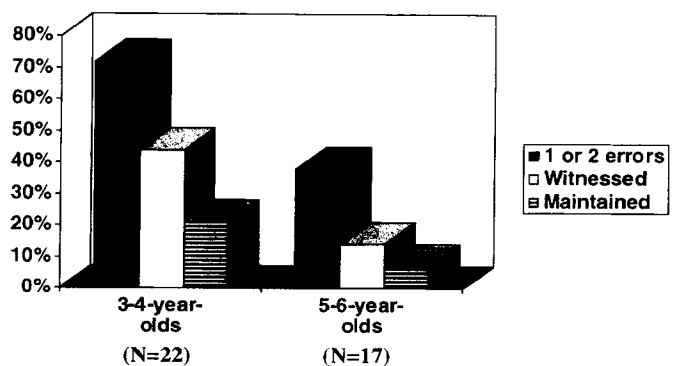


Figure 3. Percentage of preschoolers' answers that were erroneous. Light-colored bar indicates that the child asserted that an incorrect event occurred; dark-colored bar indicates that the child claimed to have actually observed the nonevent; hash-marked bar indicates that the child insisted on having witnessed the event, despite mild attempt at dissuading.

### SUGGESTION GROUP (NO STEREOTYPE; SUGGESTIONS)

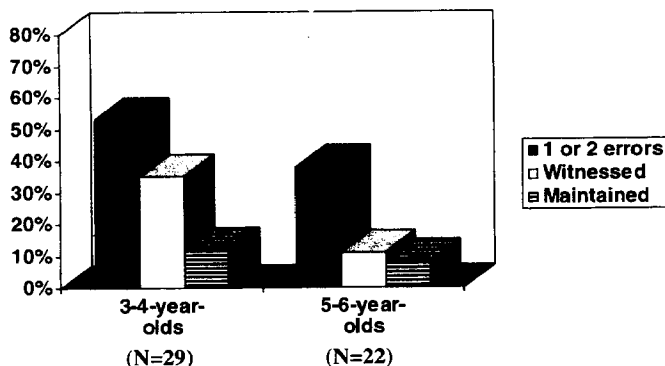


Figure 4. Percentage of preschoolers' answers that were erroneous. Light-colored bar indicates that the child asserted that an incorrect event occurred; dark-colored bar indicates that the child claimed to have actually observed the nonevent; hash-marked bar indicates that the child insisted on having witnessed the event, despite mild attempt at dissuading.

ously reported in their free narratives that Sam Stone had carried out one or both misdeeds. Nothing approaching this level of suggestibility has heretofore been reported in the memory development literature, a function no doubt of the present study's use of repeated suggestions combined with a set of congruent expectancies. It will be important to replicate and extend this finding, given its rarity.

In response to follow-up probes, 72% of the youngest preschoolers' responses indicated that Sam Stone did one or both misdeeds, a figure that dropped to 44% when asked if they actually saw him do these things. It is important to note that 21% continued to insist that they saw him do these things, even when gently challenged with a countersuggestion. For the older preschoolers, the situation, though better at all levels, was still cause for concern, as seen in Figure 4.

In view of the data described earlier, we can ask to what extent the four groups of preschoolers differentially reported what happened during Sam Stone's visit. To address this question, we conducted a 2 (age)  $\times$  4 (group) repeated measures multivariate analysis of variance (MANOVA) on the "commission errors" (i.e., claiming to have witnessed a nonevent) in the free narrative, probed recall, and challenged (countersuggestion) recall data. Errors of omission, that is, failing to report a real event, were not analyzed in these models, as they could occur only in free narratives, given the erroneous nature of both probes. (Because of an absence of commission errors for the older children in challenged reports in the control condition, we analyzed each level separately.) This analysis revealed significant main effects for both age,  $F(1, 146) = 76.55, p < .0001$ , and group,  $F(3, 146) = 29.65, p < .001$ . Follow-up tests indicated that, overall, older preschoolers reported more accurately than did younger children ( $M_s = 72\%$  and  $86\%$ , for the younger and older preschoolers, respectively), and the control group reported more accurately ( $96\%$ ) than the stereotype group ( $83\%$ ), which

reported more accurately than the suggestion group ( $72\%$ ), which in turn reported more accurately than the suggestion-plus-stereotype group ( $64\%$ ), all  $p_s < .05$ . These results were qualified by a marginally reliable Age  $\times$  Group interaction, resulting from a somewhat steeper regression of errors on age in the suggestion than in the stereotype condition,  $F(3, 146) = 3.01, p < .07$ .<sup>2</sup> That is, younger preschoolers were disproportionately more impaired by the repeated erroneous suggestions than were the older preschoolers, whereas the two age groups were more similarly impaired by the stereotype induction ( $F < 1$ ).

As mentioned earlier, countersuggestions were asked only if the child assented to a false probe, otherwise there was no reason to ask the countersuggestion. If an event is inaccurately reported in free narrative (i.e., the child spontaneously volunteers misinformation), what are the odds that it will continue to be inaccurately reported under probed and challenged recall? To explore this question, we configured the data so that each of the two events were classified as either accurate or inaccurate, and we calculated the conditional probabilities of accurate reporting, given one or two prior inaccurate reports. Younger preschoolers were more likely to make inaccurate reports if they had previously made one inaccurate report ( $.36$ ) or two inaccurate reports ( $.49$ ). For older preschoolers, the same effect was apparent ( $.23$  and  $.36$ ). One-way analyses of variance examining the differences in these conditional probabilities as a function of group and age revealed that both effects were significant: for age,  $F(1, 76) = 4.81, p = .03$ ; for group,  $F(3, 76) = 3.38, p = .02$ . The group main effect was due to all but the control group being affected by the existence of a prior erroneous report, with an inaccurate report being associated with an increased likelihood of making a subsequent inaccurate report (all  $F_s > 3.10$ , all  $p_s < .05$ ).

#### Can Adults Detect Inaccurate Reports?

It was interesting to see the number of false perceptual details that children who were assigned to the stereotype-plus-suggestion condition provided to embellish their reports of nonevents. Many of these children did not simply reply yes or no to a probe but supplied richly detailed narratives (e.g. claiming that Sam Stone took the teddy bear into a bathroom and soaked it in hot water before smearing it with a crayon). So seemingly believable were their reports that we presented videotapes of 3 of our participants (a 3-year-old, a 4-year-old, and a 5-year-old) to 119 researchers and clinicians who work in the area of children's testimonial issues—to see if they could discriminate between the erroneous reports and the accurate ones. This was done at two conferences, and in both cases the results were the same. The majority of both audiences could not reliably tell overall whether the events reported by the children had occurred or not, nor could they identify which children were on the whole most accurate (see Figure 5).

The complete videos of 3 preschoolers assigned to the stereo-

<sup>2</sup> In view of the categorical nature of the dependent variable (correct vs. incorrect) in the MANOVA, we carried out the same model with a logistic-link function. The results were virtually identical, except that the marginal two-way interaction now exceeded the .05 level.

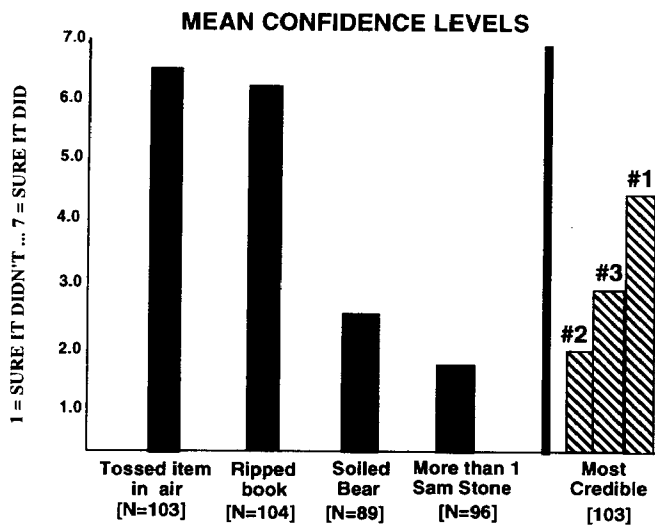


Figure 5. Mean confidence ratings of professionals viewing child interviews (1 = very confident that event did not occur, 4 = uncertain, 7 = very confident event did occur). Hash-marked bars on right side represent overall credibility ratings of 3 children, with Child 1 rated most credible and Child 2 rated least credible. In reality, Child 1 was least accurate, Child 2 was most accurate, and Child 3 was in between in accuracy.

type-plus-suggestion condition were shown to the professionals as the children gave free narratives, followed by probed and challenged recall. The particular videos were not randomly selected but were chosen because they represented the broad spectrum of answers children in this condition gave when asked about Sam Stone's visit. In addition, all of the accounts given by the 3 children were relatively coherent, and the children seemed engaged by the interviewer and confident in their answers. Child 1 was a 3-year-old girl who asserted spontaneously, and with seeming pleasure, that Sam Stone had done all of the acts listed on the left side of Figure 5 (i.e., tossed things in the air, ripped a book, soiled a bear, and been accompanied by "another Sam Stone"). Child 2 was a soft-spoken 4-year-old girl, who asserted only that Sam Stone had come into the classroom, said hello, and walked around the room before exiting, all of which was entirely accurate. At prompting, she denied knowing anything about a book or teddy bear. Child 3 was a 5-year-old boy who, while initially asserting that Sam Stone had only come in and looked around the classroom, answered the prompting questions by asserting that Sam Stone had ripped the page of a book and had painted ice cream all over a teddy bear in the school yard with a paint brush that was handy.

Audience members were told only that the study was about the visit of a person named Sam Stone to the children's classroom, and that all of the children in the videos had witnessed the same visit by Sam Stone. They were then told that they should decide for themselves what occurred during this visit, based on viewing the videotapes of interviews with 3 children about the event. Immediately after viewing the 3 children, the audience was asked to rate the accuracy of particular statements made by the children about Sam Stone's visit on a 7-point scale, where 1 = very confident that the event did not occur and 7 =

very confident that the event did occur. At no time was the impression conveyed that some of the particular events we were asking about did or did not occur; the audience was simply asked to decide in the case of each event whether they believed that it had occurred or not. Audience members were instructed to make their ratings individually, without discussing their answers with others seated nearby.

An interesting aspect of Figure 5 is that both groups of audiences rated Child 2, who was completely accurate in her account, as the least credible of the 3 children. Conversely, the audience rated Child 1, whose account contained by far the most inaccurate assertions about Sam Stone's visit, as the most credible. Similarly, for the four specific events that were addressed by the 3 children, both audiences were unable to reliably identify whether the event had occurred; they were better than chance at one of the four specific questions (i.e., they rightly determined that there was not "more than one Sam Stone"), they were exactly at chance on one of the four questions (i.e., they were undecided about whether Sam Stone had soiled the bear), and they were reliably below chance on the remaining two questions (i.e., they felt fairly certain that Sam Stone had tossed an item in the air and had ripped a book during his visit).

An additional method of assessing these professionals' ratings was carried out. Hits and false alarms were tallied for each professional, summed across the 3 children and the four specific events. Decision matrices were constructed for these aggregated data, with hits along the y-axis and false alarms along the x-axis (Banks, 1970), allowing a Receiver Operator Characteristics curve to be plotted. The area under the curve ( $A'$ ) was .52, corresponding to a  $d'$  near zero (i.e., chance). This is persuasive evidence that professionals cannot reliably discriminate signals from noise in the determination of children's statements if the children have been persistently rehearsed and provided with congruent stereotypes.

This procedure of questioning audiences about Sam Stone's visit on the basis of the videotaped interviews of 3 children was clearly not a systematic study of the factors that influence adults' beliefs about children's accuracy. Rather, our goal was to see whether, on the face of it, audiences of professionals would be able to consistently "see through" the inaccuracies (in this case, the commission errors) of the children's reports to arrive at the truth about a past occurrence. We added this feature to the study to test the opinion sometimes expressed that young children's erroneous reports can be easily detected by adults (e.g., Goodman, 1990). Thus, we were interested in seeing whether this would be the case for children who had been repeatedly suggestively questioned about nonevents over long intervals. Because this audience-rating procedure did not allow us to systematically covary (between children) elements of the reports that may have affected professionals' perceptions of the truth, we cannot be sure about the specific factors that influenced the audiences' perceptions.<sup>3</sup>

<sup>3</sup> Although we do not claim to have provided irrefutable evidence that professionals cannot reliably discriminate between accurate and false reports of children who have been subjected to persistent suggestive questioning over long delays, we do think that the evidence, such as it is, accords with this conclusion. We base this assertion on two additional

Nonetheless, we can say anecdotally that all of the children we showed to audiences appeared to us to be fairly comfortable, confident, and adept at relaying their stories. Because Child 1 indicated that more events occurred during Sam Stone's visit, her report was naturally longer than the other 2 children's reports (3 min 30 s), which were both under 2 min in length (Child 2's report was 1 min 10 s; Child 3's was 1 min 50 s). Child 1 also provided the most perceptual detail, Child 3 the second most, and Child 2 the least, which was reasonable because Child 2's story contained little in the way of activity that would invite such detail (i.e., quite accurately, she asserted that Sam Stone had come into the classroom and said hello and looked around, but that "nothing happened"). The amount of perceptual detail, then, appears to have been in accord with raters' conclusions about the child's credibility (the more detail, the more credible the child). However, it is interesting to note that this may not account for the entire heuristic used by our audience members to judge the accuracy of particular events. To wit, Child 1, the most credible in the eyes of our audience, was nonetheless the one who asserted that there was "more than one Sam Stone," which was the only one of the four fabricated events that audiences were reliably able to determine had not occurred.

The problems audiences had in determining that some of the events did not occur may not have been one of mistrusting Child 2, although she was the most reticent of the children. Rather, it seems that Child 1 and Child 3 were so compelling in their accounts that audience members were "taken in" by at least the gist of their reports. In the case of the assertion that there was "more than one Sam Stone," audiences may have correctly doubted this because it came at the end of a series of assertions about what a single Sam Stone had done, and only one Sam Stone was referred to by the other 2 children. Hence, at

---

pieces of evidence: First, we showed these same three videotapes to over 1,500 psychiatrists, social workers, and clinical psychologists at several professional meetings during 1994–1995. The results were exactly the same; these mental health professionals, many of whom testify in court as "validators" of child abuse, were no better than chance at detecting accurate reports. In addition, we showed videotapes of 10 children from another study in which they were repeatedly interviewed about non-events over a similarly long delay (12 weeks) to professionals and asked them to do the same 7-point rating procedure used in this study (Ceci, Loftus, Leichtman, & Bruck, 1994). Again, the results were the same. Professionals were no better than chance at distinguishing between the children's accurate and inaccurate narratives: There were as many professionals who were reliably worse than chance at detecting which events were real as there were professionals above chance (overall  $p = .60$ , for two-tailed test,  $\alpha = .025$  for each tail). A static Bernoulli sampling process specifies the likelihood of correctly judging a real claim ( $p$ ) and the likelihood of achieving precisely  $x$  correct in  $N$  independent trials =  $(N/x) p^x q^{N-x}$ , where the probabilities for  $x = 0-10$  correct guesses,  $N = 10$  trials,  $p = .5$ , and  $q = .5$ . (A two-tailed test was preferred in view of our interest in the number of raters who performed above as well as below chance.) Finally, our finding accords with Horner, Guyer, and Kalter's (1993) results using a different methodology. In their study, clinical psychologists' and social workers' predictions of the accuracy of children's sexual abuse reports were disturbingly unreliable, spanning the full range of estimated probabilities (from 0 to 1.0) of the child having been abused.

least part of the reason for adult skepticism may have been rooted in the coherence of report content, and not in the manner in which the information was conveyed.

It is not that the members of these audiences were worse than anyone else at assessing which children gave accurate accounts, but that the accuracy of children's reports is extremely difficult to discern when children have been subjected to repeated erroneous suggestions over long retention intervals, especially when coupled with the induction of stereotypes. These findings do not support the claims of those who think that it is easy to detect a young child's false report and lend at least anecdotal support to recent conclusions that such a task may be quite difficult. The reason is that, unlike in the modal study in which a child is presented a single erroneous suggestion, these children received persistent and intensive suggestions.

It may be that the children who received repeated false suggestions actually incorporated the erroneous suggestions into their memories. We cannot tell from these data, as the experiment was not designed to separate this from the alternative possibility of nonmemorial distortion. But the children's reports certainly seemed compelling to the professionals to whom we showed them.

## Conclusion

In conclusion, the results of this study extend our knowledge about children's report accuracy by adding two factors that have been neglected in past research, namely, the provision of a stereotype and the use of repeated erroneous suggestions over a long delay period. Although of the all children suffered from a combination of these factors, the deleterious effect was particularly pronounced for the youngest preschoolers. The reason for their disproportionate susceptibility is unclear, but two candidates can be mentioned. First, the repeated presence of suggestions of the type that were used in this study (e.g., "When Sam Stone soiled the teddy bear, did he do it on purpose or was it an accident?") may have induced children to form mental images of these nonevents. If so, then the request for a free narrative during the fifth (forensic) interview may have created a source-separation challenge. Younger children may have fared more poorly on our tasks because of source-monitoring problems; they are known to have greater difficulty than older children in determining whether mental images are the result of past experiences in the world or are products of their past imaginings. Alternatively, younger children may have simply been more sensitive to the demand characteristics of the situations constructed at the time that the stereotype was laid down and that leading questions were posed. These 3- and 4-year-olds may have been aware at the time of the final interview that their own memories of Sam Stone's visit were in conflict with the suggestions they had received from adults about his potential behavior, and the developmental trend obtained may be the result of their trusting an adult's interpretation more than their own. On the basis of our results, a combination of these two alternative explanations may also be possible. Namely, it may be the case that the original reason for children's acquiescence to suggestions during the four initial interviews was a social one. However, once these children had committed to the idea of misdeeds conducted by Sam Stone in their school, the cognitive mechanism just de-



scribed may have kicked in, and source-monitoring problems may have arisen at some unspecified point during the course of the period between the original event and the free narrative recounting. As the youngest children discussed and thought about Sam Stone's clumsy behavior during multiple interview sessions, the strength of their memory traces for these events would have been likely to increase, concomitantly increasing the probability of source misattributions at a later point. Although 5- and 6-year olds, and for that matter adults in analogous situations, do not have immunity to the social and memorial factors that potentiate report distortion, our older participants' greater ability to both resist situational demand characteristics and to accurately monitor their memory sources may have combined to render them dramatically less vulnerable to suggestion than their younger counterparts.

Given the varied results that emerged from the four conditions in this experiment, it is clear that children's suggestibility is best viewed as heavily reliant on the entire context in which event reporting takes place. This context includes the cognitive framework that is set up before the memory trace is laid down (e.g., including beliefs and stereotypes that relate to the encoded information) and the nature of information pertinent to the event of interest that is encoded after that event has occurred (e.g., information suggested during storage). Given this picture of the multiple points at which misleading information may cause report distortion, it behooves us to consider suggestibility throughout development as a statelike, as opposed to a traitlike, quality. Although younger children show a greater vulnerability to both preevent and postevent suggestions, it is evident from the data we have presented that situations may be engineered in which even very young children's reports are wholly accurate. As demonstrated by our control group, when the context of a child's reporting of an event is free of the strong stereotypes and repeated leading questions that may be introduced by adults, the odds are tilted in favor of factual reporting.

## References

- Banks, W. P. (1970). Signal detection theory and human memory. *Psychological Bulletin*, 74, 81-89.
- Brainerd, C. J., & Reyna, V. (1988). Memory loci of suggestibility development: Comment on Ceci, Ross, and Toglia. *Journal of Experimental Psychology: General*, 118, 197-200.
- Brainerd, C. J., Reyna, V. F., Howe, M. L., & Kingma, J. (1990). The development of forgetting and reminiscence. *Monographs of the Society for Research in Child Development*, 55(3-4, Serial No. 222).
- Bussey, K. (1992). Children's lying and truthfulness: Implications for children's testimony. In S. J. Ceci, M. Leichtman, & M. Putnick (Eds.), *Cognitive and social factors in preschoolers' deception* (pp. 89-110). Hillsdale, NJ: Erlbaum.
- Ceci, S. J., & Bruck, M. (1993). Suggestibility of the child witness: A historical review and synthesis. *Psychological Bulletin*, 113, 403-439.
- Ceci, S. J., Leichtman, M. D., & Bruck, M. (in press). The suggestibility of children's eyewitness reports: Methodological issues. In F. Weinert & W. Schneider (Eds.), *Memory development: State of the art and future directions*. Hillsdale, NJ: Erlbaum.
- Ceci, S. J., Leichtman, M., & Putnick, M. (Eds.). (1992). *Cognitive and social factors in early deception*. Hillsdale, NJ: Erlbaum.
- 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.
- Ceci, S. J., Ross, D., & Toglia, M. (1987). Age differences in suggestibility: Psychological implications. *Journal of Experimental Psychology: General*, 117, 38-49.
- Ceci, S. J., Toglia, M., & Ross, D., (1988). On remembering . . . more or less. *Journal of Experimental Psychology: General*, 118, 250-262.
- Foley, M. A., & Johnson, M. K. (1985). Confusions between memories for performed and imagined actions. *Child Development*, 56, 1145-1155.
- Foley, M. A., Johnson, M. K., & Raye, C. L. (1983). Age-related confusion between memories for thoughts and memories for speech. *Child Development*, 54, 51-60.
- Goodman, G. S. (1990, August). Media effects and children's testimony. In D. Singer (Chair), *The impact of the media on the judicial system*. Symposium conducted at the 98th Annual Convention of the American Psychological Association, Boston.
- Goodman, G. S., & Clarke-Stewart, A. (1991). Suggestibility in children's testimony: Implications for child sexual abuse investigations. In J. L. Doris (Ed.), *The suggestibility of children's recollections* (pp. 92-105). Washington, DC: American Psychological Association.
- Goodman, G. S., Taub, E. P., Jones, D., England, P., Port, L., Rudy, L., & Prado-Esrada, L. (1992). The emotional effects of criminal court testimony on child sexual assault victims. *Monograph of the Society for Research in Child Development*, 57(Serial No. 229), 1-152.
- Gray, E. (1993). *Unequal justice: The prosecution of child sexual abuse*. New York: MacMillan.
- Horner, T. M., Guyer, M. J., & Kalter, N. M. (1993). Clinical expertise and the assessment of child sexual abuse. *Journal of the American Academy of Child and Adolescent Psychiatry*, 32, 925-931.
- Howe, M. L. (1991). Misleading children's story recall: Forgetting and reminiscence of the facts. *Developmental Psychology*, 27, 746-762.
- Humphrey, H. H., III (1985). *Report on Scott County investigations*. Minneapolis, MN: Attorney General's Office.
- King, M., & Yuille, J. (1987). Suggestibility of the child witness. In S. J. Ceci, D. Ross, & M. Toglia (Eds.), *Children's eyewitness memory* (pp. 24-35). New York: Springer-Verlag.
- Lepore, S. J., & Sescio, B. (1994). Distorting children's reports and interpretations of events through suggestion. *Applied Psychology*, 79, 108-120.
- Lewis, M., Stranger, C., & Sullivan, M. (1989). Deception in three-year-olds. *Developmental Psychology*, 25, 439-443.
- Lindsay, D. S., Johnson, M. K., & Kwon, P. (1991). Developmental changes in memory source monitoring. *Journal of Experimental Child Psychology*, 52, 297-318.
- Martin, C. L., & Halverson, C. F. (1983). The effects of sex-typing schemas on young children's memory. *Child Development*, 54, 563-574.
- McGough, L. (1994). *Child witnesses: Fragile voices in the American legal system*. New Haven, CT: Yale University Press.
- People v. Buckley, No. A-75-0900 (Cal. Crim. Dist. Ct. 1984).
- Peters, D. P. (1991, April). Confrontational stress and children's testimony. In M. DeSimone & M. Toglia (Chairs), *Lying and truthfulness among young children: Implications for their participation in legal proceedings*. Symposium conducted at the biennial meeting of the Society for Research in Child Development, Seattle, WA.
- Poole, D. A., & White, L. T. (1991). Effects of question repetition on the eyewitness testimony of children and adults. *Developmental Psychology*, 27, 975-986.
- Poole, D. A., & White, L. T. (1993). Two years later: Effects of question repetition and retention interval on the eyewitness testimony of children and adults. *Developmental Psychology*, 29, 844-853.
- Strangor, C., & McMillan, D. (1992). Memory for expectancy-congruent and expectancy-incongruent information: A review of the social and social developmental literatures. *Psychological Bulletin*, 111, 42-61.
- Texas v. Macias, 733 S.W.2d 192 (Tex. Crim. App. 1987).

Warren, A. R., Hulse-Trotter, K., & Tubbs, E. (1991). Inducing resistance to suggestibility in children. *Law and Human Behavior, 15*, 273-285.

Zaragoza, M., Dahlgren, D., & Muench, J. (1992). The role of memory

impairment in children's suggestibility. In M. L. Howe, C. J. Brainerd, & V. F. Reyna (Eds.), *Development of long term retention* (pp. 184-216). New York: Springer-Verlag.

## Appendix A

### Scenarios Provided to Children in the Stereotype and Stereotype-Plus-Suggestion Conditions

(Below are abbreviated versions of scripts, presented with appropriate elaboration.)

1. Sam Stone borrowed my favorite pen. Then he lost it. (But he bought me a new one.)
2. Sam Stone accidentally spilled his glass of Coca-Cola all over me. (But he cleaned it up.)
3. Sam Stone was playing with my Barbie doll, and he tripped and dropped her. (But he got her fixed.)
4. Sam Stone was using a tape recorder, and he accidentally broke it. (But he got it fixed.)
5. Sam Stone wanted to take a picture of me with my camera, but as he went to do so, he accidentally dropped the camera. (But he got the camera fixed.)
6. Sam Stone asked if he could have a lick of my ice-cream cone. But when he went to take a lick, the cone fell on the floor. (But he got me a new ice-cream cone.)
7. Sam Stone was trying on my jacket, and he accidentally ripped it because it was too small for him. (But he got the hole sewn for me.)
8. After playing "Candyland" at my house, Sam Stone took some of the pieces of the game home with him in his pockets. He didn't

realize that he had the pieces in his pockets until the next day when I phoned him. (When he discovered the pieces, he brought them back.)

9. Sam Stone visited my house, where there were some flowers in a vase. When he bent down to smell the flowers, the vase was knocked over, and the flowers and water spilled all over the floor. (But he didn't mean to do it, and helped clean up the spill.)
10. Sam Stone came to visit on a very rainy day, and he had mud all over his shoes. He said that he would take off his shoes before coming into the house, but then he forgot; he got mud all over the kitchen. (But when he realized what he'd done, he wiped up all of the mud.)
11. Sam Stone and I were playing in a sandbox, and he was eating a lollipop. But while he was playing, the lollipop slipped, and it fell in the sand. There was so much sand stuck to the lollipop that it couldn't even be washed off. (Sam had to get a whole new lollipop.)
12. When Sam Stone was painting at school, some children went to look at his picture, but he knocked a can of red paint all over the floor and the children. (But he was very sorry and helped clean up the paint.)

## Appendix B

### Interview Protocols and Leading Questions Asked of Children in the Suggestion and Stereotype-Plus-Suggestion Conditions

#### Interview 1

With the teachers' cooperation, children were taken out of the classroom individually by the interviewer to go and see a toy airplane that twirled around on a stick. Once alone with the interviewer (in a quiet space very near the classroom), each child was engaged in a conversation about the airplane and was allowed to try it out. After a few moments of play, the plane was set aside and the interviewer showed the child a book that the teacher had read to the class the day before. The interviewer leafed through the book and noticed a torn page. The child was asked the following questions:

1. "Who ripped the book?" (If child responded "don't know," he or she was asked: "Who do you think might have ripped the book?")

The child was then shown a red-and-white bear that had been stained. The following questions were then asked regarding the bear.

2. "Who got the bear dirty?" (If child responded "don't know," he/she was asked: "Who do you think could have gotten the bear dirty?")

In closing the interview, the interviewer engaged the child in a discussion about a new Barbie (for girls) or action figure (for boys) doll that she had just bought and had not yet named. The interviewer promised that she would bring the Barbie or action figure doll the following

week for the child to see. She then brought the child back to the classroom.

#### Interview 2

Children were taken out of the classroom individually by the interviewer, with the enticement of seeing the Barbie doll or action figure discussed in the first interview. After allowing the child to see the doll and conversing about possible names for it, the interviewer showed the child the book that had been discussed during the first interview. The following questions were then posed:

1. "Remember when Sam Stone ripped the book? Did he rip it on purpose, or by accident?"
2. "Did Sam Stone rip the book with his hands, or did he use scissors?"
3. "When Sam Stone ripped the book, was he in the classroom, in the hallway, or in the bathroom?"

The soiled red-and-white bear was then reintroduced, and the following questions were asked:

4. "When Sam Stone got the bear dirty, did he do it by accident, or on purpose?"
5. "Did Sam Stone get the bear dirty with chocolate ice-cream, chocolate syrup, Hershey kisses, crayons, or something else?"

6. "When Sam Stone got the bear dirty, was he in the classroom, the hallway, or the bathroom?"

At the completion of questioning, the interviewer brought the child back to the classroom.

### Interview 3

The child was taken out of the classroom in order to see an attractive pop-up toy that the researcher had in an adjacent room. After allowing the child to play with the toy, the researcher asked the following questions (without props):

1. "When Sam Stone ripped the book, did he do it because he was angry, or by mistake?"
2. "When Sam Stone ripped the book, was it before or after naptime?"
3. "Was Sam Stone happy or sad that he got the bear dirty?"
4. "When Sam Stone got the bear dirty, was he wearing long or short pants?"

On completing the questioning, the interviewer brought the child back to the classroom.

### Interview 4

The child was taken out of the classroom to "talk to" the interviewer "for a minute." (The interviewer was now well-known to the children.) The interviewer opened the interview by saying, "I just wanted to ask you some things about that day that Sam Stone visited your classroom," and then proceeded with the following questions:

1. "When Sam Stone ripped the book, did he do it alone or with a helper?"
2. "When Sam Stone ripped the book, did he tell the teacher that he did it?"
3. "What did the teacher say when she found out that he ripped it?"
4. "When Sam Stone got the bear dirty, what was that stuff he got on it?"
5. "Did he bring that stuff from home, or did he get it at school?"
6. "Where did he hide that stuff; in his pocket, or in a bag?"
7. "When Sam Stone got the bear dirty, did he tell the teacher that he did it?"
8. "What did the teacher say when she found out that he got the bear dirty?"

## Appendix C

### Protocol and Questions Asked of All Children in the Final (Fifth) Interview

With the cooperation of their teachers, children were taken out of the classroom individually by the interviewer, whom they were told wanted to ask them "some questions." Children chatted with the interviewer on the way to the interview room, and for up to several minutes once in the room, about pleasantries. Once in the room, children were seated comfortably on an upholstered chair facing the interviewer. The interviewer gained the child's attention and told him or her, "I have an important question for you." The following "free narrative" question was then asked:

1. "Remember that day that Sam Stone came to your classroom? Well, I wasn't there that day, and I'd like you to tell me everything that happened when he visited. Can you tell me what happened?"

Children were given time to tell as much about Sam Stone's visit as they could, and were asked, "Can you remember anything else?" followed by a chance to respond, until they indicated that they had nothing else to tell.

All children were then given the following specific "prompting" questions, unless they had referred specifically to the items mentioned in these questions in their free-narrative answers.

2. "I heard something about a book. Do you know anything about that?"
3. "I heard something about a teddy bear. Do you know anything about that?"

After both of these questions were posed, all children who indicated (either in response to the free narrative or the prompting) that Sam Stone did something to the book or the teddy bear during his visit were asked the following questions:

4. "Did you see him (action vis-a-vis the book, as noted by the child, inserted here; e.g., *rip the book*) with your own eyes?"
5. "Did you see him (action vis-a-vis the teddy bear, as noted by the child, inserted here; e.g., *put paint on the teddy bear*) with your own eyes?"

In the case of only those children who said that they saw Sam Stone commit the acts in question with their own eyes, countersuggestion questions were then posed:

6. "You didn't really see him (action vis-a-vis the book, as noted by the child, inserted here; e.g., *rip the book*), did you?"
7. "You didn't really see him (action vis-a-vis the teddy bear, as noted by the child, inserted here; e.g., *put paint on the teddy bear*), did you?"

Received September 8, 1993  
 Revision received September 30, 1994  
 Accepted September 30, 1994 ■