

May 14, 2018

63. Children's conversational memory regarding a minor transgression and a subsequent interview.

Stacia N. Stolzenberg, *Arizona State University*

Kelly McWilliams, *CUNY John Jay College*

Thomas D. Lyon

Children's Conversational Memory Regarding a Minor Transgression and a Subsequent Interview

Stacia N. Stolzenberg
Arizona State University

Kelly McWilliams
City University of New York

Thomas D. Lyon
University of Southern California

Children's memories for their conversations are commonly explored in child abuse cases. In 2 studies, we examined conversational recall in 154 4- to 9-year-old children's reports of an interaction with a stranger, some of whom were complicit in a transgression and were admonished to keep it a secret. Immediately afterward, all children were interviewed about their interaction. One week later, children were asked recall questions about their interaction with the stranger, their conversations with the stranger, and their conversations with the interviewer. Overall, interaction recall questions elicited few details about children's conversations, whereas conversation recall questions were effective in doing so. Accuracy was high in response to both the interaction and conversation recall questions, with no differences observed. Questions explicitly inquiring about coaching elicited higher error rates, as well as apparent attempts to maintain secrecy. Source errors were rare. Conversation recall questions elicited new transgression disclosures among a substantial percentage of children. The results provide tentative support for the use of recall questions in eliciting conversational information from children.

Keywords: conversational memory, child interviewing, children's memory, conversational recall

Researchers have called the study of memories for conversations the "orphan child" of research on witness performance (Davis & Friedman, 2007). Apart from the classic finding that the exact words one hears are quickly forgotten and replaced with gist (Sachs, 1967), a finding subsequently extended to difficulty remembering the verbatim statements of prior conversations (Stafford & Daly, 1984), there is little researchers can say about memory for conversations. Even less attention has been paid to children's memory for conversations, despite an explosion of interest in children's capacities as witnesses and frequent questioning of children in court regarding their conversations about alleged abuse (Lyon & Stolzenberg, 2014).

We conducted two studies examining 4- to 9-year-old children's gist recall of conversations with a stranger and an interviewer about a potentially negative event involving the stranger. In what

follows, we first discuss the legal significance of child witnesses' testimony about conversations, noting how relevant conversations are admissible in criminal trials, particularly in child sexual abuse cases. We then discuss how practice guides advise child interviewers to ask children about conversations and emphasize the recommendations that they attempt to elicit as much information as possible through recall questions. Finally, we discuss the limited research on children's memory for conversations, both the observational research examining how alleged child sexual abuse victims are questioned in the field, and laboratory research examining children's memory for conversations.

The Legal Significance of Children's Testimony About Conversations

Most testimony from children in criminal court concerns allegations of sexual abuse, physical abuse, and exposure to domestic violence (Goodman, Quas, Bulkley, & Shapiro, 1999). Because the focus in these cases is whether the acts occurred, rather than what was said, one might suppose that when children testify about abuse, their prior conversations are of limited relevance or constitute inadmissible hearsay. However, in a study examining criminal child sexual abuse trials in a California sample, child witnesses were asked about conversations with disclosure recipients in 100% of the cases and were asked about conversations with suspects in 92% of the cases (Stolzenberg & Lyon, 2014).

In the United States, children's disclosures prior to their testimony are often legally admissible to prove that abuse occurred under exceptions to the rule against hearsay. The exceptions include spontaneous statements (also known as excited utterances),

Stacia N. Stolzenberg, School of Criminology and Criminal Justice, Arizona State University; Kelly McWilliams, Department of Psychology, John Jay School of Criminal Justice, City University of New York; Thomas D. Lyon, Gould School of Law, University of Southern California.

This research was supported by NICHD grants HD047290 and HD087685 to Thomas D. Lyon. Portions of this data were presented at the annual meeting of the International Investigative Interviewing Research Group in 2014, and at the American Psychology-Law Society in 2015 and 2018.

Correspondence concerning this article should be addressed to Stacia N. Stolzenberg, School of Criminology and Criminal Justice, Arizona State University, 411 North Central Avenue, Suite 600, Phoenix, AZ 85004. E-mail: stacia.stolzenberg@asu.edu

statements made for the purpose of medical diagnosis and treatment, prior consistent statements, and statements by a party-opponent (i.e., the defendant when the statement is offered by the prosecution; Duke, Lee, & Payer, 2007; Myers, 2016). Furthermore, adults' statements to children are often admissible as non-hearsay, insofar as they are not "offered to prove the truth of the matter asserted" but to prove that the statement was made (Mueller & Kirkpatrick, 2015). For example, this would include disclosure recipients' statements to children offered to prove that they put pressure on the child, either to make a false disclosure or a false denial (Lyon & Stolzenberg, 2014). Also typically admissible are suspects' statements to children offered to demonstrate the means by which the suspects may have secured children's acquiescence to abuse (Leclerc, Proulx, & Beaugard, 2009) or guaranteed children's silence (Smith & Elstein, 1993).

Questioning Child Witnesses About Conversations: Question-Type

Practice guides for interviewers questioning children about abuse emphasize the need to maximize the use of recall questions because this minimizes the likelihood of commission errors and because recall questions produce more details than recognition questions (Lamb, Hershkowitz, Orbach, & Esplin, 2008). Conversely, recognition questions increase commission errors if they contain inaccurate information, and such errors are most likely when questions are overtly leading, such as tag questions (e.g., "Your mother told you what to say, didn't she?"; Cassel, Roebbers, & Bjorklund, 1996; Greenstock & Pipe, 1996).

Ideally, children questioned about abuse would be questioned with nothing but free recall questions, using "Tell me why you came to see me" to elicit the allegation, and "Tell me everything that happened," "What happened next," and "Tell me more about [detail]" questions, also known as invitations, to elicit the abuse report. However, more specific questions may be necessary to elicit details from younger children (Hershkowitz, Lamb, Orbach, Katz, & Horowitz, 2012), from children who are reluctant to respond (Hershkowitz & Elul, 1999), and to elicit specific details that children may not mention spontaneously, including information about abuse disclosure and perpetrator statements.

The National Institute of Child Health and Human Development (NICHD) Investigative Interviewing Protocol is the most-researched protocol for interviewing children about abuse. In the latest version of the protocol (*Revised Investigative Interview Protocol, 2014*), conversation questions are recommended in two places. First, if the child fails to disclose in response to less direct inquiries, one of the recommended questions is "I've heard that you talked to [a doctor, teacher, social worker, other professional] at [time, location]. Please tell me what you talked about." Similarly, the American Professional Society on the Abuse of Children (APSAC) guidelines state that "prompts referencing the person to whom the disclosure was made. . . .can be useful" in eliciting a disclosure (APSAC, 2012, p. 18). Notably, however, the purpose of these questions is to elicit the abuse report, rather than the conversation itself. Second, during the "disclosure" phase of the interview, the NICHD protocol recommends that the interviewer ask "Now I want to understand how other people found out about [the allegation]" and "Does anybody else know what happened?" (*Revised Investigative Interview Protocol, 2014*).

Lyon and Stolzenberg (2014) have argued that children could profitably be asked more specific recall questions about conversations with disclosure recipients, abuse suspects, and other influential adults (such as parents who might coach the child). These questions would go beyond the NICHD protocol and ask about specific interlocutors, including "What did you say to [disclosure recipient]" "What did [disclosure recipient] say," and "What has [the suspect] said to you about [the abuse]" (Ahern & Lyon, 2011).

Observational Research Examining Children's Conversational Memory

Observational research examining abuse interviews suggests that children may need to be specifically asked about their conversations in order to elicit conversational memories. Malloy, Brubacher, and Lamb (2013) examined interviews of 204 5- to 13-year-olds using the NICHD protocol about sexual abuse. Forty percent of the children had not mentioned a disclosure recipient prior to questions explicitly inquiring into whether someone knew of the abuse. Ahern and Lamb (2017) examined 95 interviews conducted by law enforcement either using the Memorandum of Good Practice (MoGP; Home Office, 1992) or the NICHD protocol. The MoGP recommends asking about disclosure but does not provide any guidance regarding how to ask. The authors found that 55% of the MoGP interviews failed to inquire into disclosure. The NICHD interviews elicited more information about disclosure, leading the authors to posit that "interviewers may often need to ask for this information explicitly" (p. 7). Surprisingly, no observational research has examined forensic interviewers' questions about conversations with abuse suspects.

Laboratory Research Examining Children's Conversational Memory

Although child witness researchers have not focused on conversational memory, they have included questions about conversations in techniques designed to enhance children's incomplete recall. Saywitz and colleagues developed Narrative Elaboration Training (NET), in which children are trained to provide more complete narratives through the use of cue cards depicting different aspects of events, including a card depicting a person with a talk bubble that is described as reflecting what people said and thought (Saywitz & Snyder, 1996). Children from 3 to 11 years old have been found to produce more information about experienced events after being trained with the cards, without a decrease in accuracy (Bowen & Howie, 2002; Brown & Pipe, 2003a, 2003b; Camparo, Wagner, & Saywitz, 2001; Dorado & Saywitz, 2001; Nathanson, Crank, Saywitz, & Ruegg, 2007; Peterson, Warren, & Hayes, 2013; Saywitz & Snyder, 1996; Saywitz, Snyder, & Lamplear, 1996). Subsequent work has found that use of the recall questions alone (e.g., "Tell me more about what people said and thought") increases recall without decreasing accuracy (Brown & Pipe, 2003a; Kulkofsky, 2010). Across studies, recall accuracies tended to be very high, including questions about conversations; for example, Brown and Pipe (2003b) found reports of what people said and thought to be about 90% accurate for 7- to 8-year-olds. Of course, given the speed with which verbatim memory for conversations is lost, even in adults, the research has always examined memory for gist.

However, the work on NET and recall questions is limited with respect to conversational memory. First, most of the research does not analyze conversational recall memory apart from other aspects of children's event memory (Dorado & Saywitz, 2001; Kulkofsky, 2010; Nathanson et al., 2007; Peterson et al., 2013). When it does, NET and recall questions sometimes have been found to increase conversational content (Brown & Pipe, 2003a; Saywitz et al., 1996; Saywitz & Snyder, 1996), but not always (Bowen & Howie, 2002; Brown & Pipe, 2003b; Camparo et al., 2001). Second, asking about what "people" said may be insufficiently specific with respect to speaker to be useful; as noted earlier, a key question is often who said what to whom. Third, as Kulkofsky (2010) noted, increasing recall without decreasing accuracy often means that both the number of accurate and inaccurate details increase; that is, the raw numbers of errors increase, but the proportion of errors stays the same. Even if the percentage accuracy is unaffected, it is important to conduct a qualitative assessment of the errors to assure that the increased recall is worthwhile.

Lawson and London (2015) examined 8-year-olds conversational gist recall memory after a 1- or 3-week delay and reported a 32% error rate (Lawson & London, 2015). Questioning children a year later, they found similarly poor recall (Lawson & London, 2017). The conversation in Lawson and London was a story children made up with the encouragement of an adult interlocutor. Lawson and London (2015) suggested two reasons for the unusually high rate of error. First, they noted that "children might have reported inaccurate information in efforts to tell a good story rather than to recall the target conversation" (p. 441). This highlights a distinction between retelling and recalling; memory researchers have noted how norms of retelling encourage embellishment (Marsh, 2007). Indeed, 67% of children's errors in Lawson and London (2015) were elaborations, in which they added details, and another 8% were inferences. Second, Lawson and London (2015) noted that when children recall conversations about events, as opposed to recalling a story, their memory for conversations might be exaggerated by their memory for the events themselves.

Given the type of conversations children are asked about in court (Stolzenberg & Lyon, 2014), it is important to further assess children's recall of conversations about events. On the one hand, children's recall for an event might facilitate their recall for conversations about the event, much like script recall can facilitate recall of individual instantiations of a script-like event (Brubacher, Roberts, & Powell, 2012; Connolly & Gordon, 2014). On the other hand, event memories might interfere with conversational memories, creating difficulties in source monitoring. That is, children could confuse their memory for an event with their memory for conversations about the event, and thus erroneously report that they *discussed* information about the event that they only *experienced*. This confusion is related to the classic finding in child suggestibility research that young children may erroneously report that they *experienced* events that they only heard *discussed* (e.g., Principe & Schindewolf, 2012).

Source monitoring difficulties may lead to other errors in recalling conversations as well. In a review of research on conversational memory in adults, Davis, Kemmelmeier, and Follette (2005) noted that a number of different types of source confusion can occur when distinguishing among speakers, distinguishing among conversations, and distinguishing what was thought from what was said. Children exhibit large developmental changes in

source monitoring abilities (Foley, 2014), though some distinctions take longer to develop than others. For example, Foley, Johnson, and Raye (1983) found developmental differences in 6- to 10-year-olds' ability to distinguish between words they imagined saying and words they said, but not in children's ability to distinguish which speakers said which words. They argued that younger children might have special difficulty with realization judgments, either because their memories for their thoughts and utterances were more similar than older children's, or because they had yet to develop the metacognitive ability to identify the difference between the two. Hence one might expect to see younger children exhibiting the greatest difficulty in determining whether they had disclosed information or had only thought about disclosing.

Another limitation to prior research on children's conversational memory is the lack of attention to questioning children about conversations regarding events that they are motivated to conceal, of obvious importance when considering children's disclosure of abuse. Children reluctant to disclose events are likely to both fail to disclose the events and omit details about discussion of those events. Younger children's limited theory of mind and inhibitory skills might lead them to disclose an adults' admonishments to keep a transgression a secret, without recognizing that by doing so they were giving away the transgression. Young children have an incomplete understanding of the nature of secret-keeping (Peskin & Ardino, 2003) and inference (Sodian & Wimmer, 1987), such that they may not understand that disclosing secrecy inducements leads interviewers to infer that a transgression has occurred.

Study 1

This study examined children's memory for an interaction with a stranger, and with an interviewer inquiring about the interaction with the stranger. To do so, we asked recall questions about their interactions and their conversations. All of the children played with the stranger and half experienced a transgression during which toys broke in the child's hands during play; for these children, the stranger admonished the child to keep breakage a secret. The stranger had encouraged the child to play with the toys and emphasized that both might get in trouble for the breakage. All children were interviewed immediately thereafter by an interviewer about their interaction with the stranger. One week later, a second interviewer asked them to recall their interactions with the stranger and their conversations with both the stranger and the first interviewer.

We were interested in the amount and accuracy of information that children would provide about their conversations both in response to recall questions about their interactions with the stranger, and in response to recall questions specifically asking children about their conversations with the stranger and with the first interviewer. We hypothesized that the conversational questions would elicit new information, and that questions asking children about their conversations with the stranger would elicit conversation information missing when children were asked about their interactions with the stranger. We tested whether the accuracy of information children provided about their conversations was comparable to the accuracy of information children provided about their interactions. We were particularly interested in how often children would err, and what kinds of errors they would make. We

hypothesized that children would exhibit some tendency to confuse discussing play with having played. We varied whether children were asked conversation questions before or after interaction questions, to determine if being asked interaction questions affected the accuracy of children's responses to conversation questions. We hypothesized that asking children interaction questions first would facilitate their recall of conversations. Finally, we examined children's disclosure patterns across the two interviews, paying special attention to whether the conversation recall questions elicited disclosures of breakage among children who had not previously disclosed. We hypothesized that conversation questions would elicit new disclosures.

Method

The following protocol was approved by the Institutional Review Board (IRB) at the University of Southern California.

Participants. The sample included eighty-three 5- to 9-year-old children ($M = 7$ years old, $SD = 1.40$, 54% male). The sample was ethnically diverse and recruited from schools serving predominantly low-income children (61% Latino, 34% African American, 5% Biracial).

Materials and procedure. At Time 1 (T1), each child completed several tasks with a female interviewer. Afterward, the interviewer told the child that she forgot some important papers and needed to retrieve them from her office. Shortly after she left, a stranger entered the room and engaged children in toy play with six toys. He followed a script in which he demonstrated how to play with each toy and then encouraged the child to play. Children were randomly assigned to toy condition, stratified by age; approximately half of children were in the *toy break* condition ($n = 43$), whereby two of the toys broke in the child's hands, and the child was asked by the stranger not to reveal the transgression to the interviewer ("Please don't tell the lady that these toys broke. We might get in trouble if she finds out the toys broke. So [child's name] can you help me out and not tell her that the toys broke?"). In the *no break* condition ($n = 40$), the stranger and child played with toys, none broke, and the stranger did not make any statements about concealing play. The toy interaction was approximately 5 minutes in length. In both conditions, after toy play the stranger excused himself and shortly thereafter the interviewer (who was blind to whether the children were in the *toy break* or *no break* condition) returned. She conducted an interview with the child lasting approximately 10 min, first building rapport with free recall questions, akin to those in the NICHD structured interview, and then asking the child about his or her interactions with the stranger, using free recall ("Tell me everything that happened while I was gone" "What happened next?") and cued recall questions ("You said you played with the [toy], tell me everything you did with the [toy]"), also known as invitations and cued invitations. The interviewer then moved to eight recognition (yes/no) questions explicitly asking whether each toy broke; if children said "yes" they were asked "Tell me everything that happened when the [toy] broke." T1 interactions with the interviewer and stranger were recorded. More detail about the T1 methodology is presented in Stolzenberg, McWilliams, and Lyon (2017b).

One week later, at Time 2 (T2), a second interviewer questioned each child. Again, the interview began with rapport-building recall questions. She then explained to the child that she wanted to learn

about what happened the last time s/he was "here." She then asked free recall and cued recall questions about the child's interaction with the stranger, and four conversation recall questions. Approximately half of the children were asked the interaction recall questions first ($n = 46$), and half were asked the conversation recall questions first ($n = 37$). The conversation recall questions were (1) "I heard you spoke with a lady [the interviewer]. I want to know what you said to the lady. I want to know the words you used. What did you say to the lady?" (2) "Now I want to know what the lady said to you last time you were here. I want to know the words she used. What did the lady say to you?" (3) "I heard you spoke with a man [the stranger] the last time you were here. I want to know about what the man said to you. I want to know the words he used. What did he say?" and (4) "What did the man say you should say to the lady?" The first three conversation recall questions were counterbalanced for order between children; there were no differences observed by order, and as such, this is not discussed further. Each enumerated prompt was delivered as one question. The fourth question was always asked last, because it directly inquired into coaching and thus was more specific (and potentially leading) than the other questions.

Debriefing. At the end of the interview, the stranger reentered the room, and the second interviewer explained to the child that she knew the stranger would play with the child. The interviewer emphasized to all children the importance of always telling the truth.

In addition, participants in the *toy break* condition were reassured that sometimes the toys would break but that it was okay because the toys could be fixed.

Transcription and coding. Statements by the child, interviewer, and stranger in children's T1 and T2 interviews were transcribed and coded by trained research assistants. Both T1 and T2 interviews were coded for disclosure of toy play and breakage. For T2 interviews, children's responses were coded for the amount of information, newness, content, accuracy, and error type.

Amount of information. Children's responses were coded for the number of details. A detail was defined as a substantive word. For example, "We made the planes fly" would include four details (we, made, planes, fly).

Newness. Each detail was coded for whether it was new or repeated information within the interview. Analyses focused on new information (notably, only 9% of the total information was repeated).

Content. To examine content, each detail was categorized as: 1 = *conversation* (e.g., "She asked me about the toys"), 2 = *action* (e.g., "I played with the lobster"; including gestures, toy break acknowledgments, toy break denials), or 3 = *description* ("e.g., it lighted up").

Accuracy. For accuracy, each detail was categorized as: 1 = *true*, 2 = *false*, 3 = *subjective* (e.g., "I had fun playing with the toys"), or 4 = *indeterminate* (events occurring before or after the experimental interaction, such as when the interviewer walked with the child from his or her classroom to the experimental room). Conversational details did not have to be verbatim to be coded as true (e.g., "I said we played with the toys" was coded as an accurate description of "we played with the [toy] and the [toy] and the [toy]").

Error type. We examined three kinds of errors: 1 = *secrecy* (i.e., incorrectly describing inducements to secrecy), 2 = *source*

confusion (i.e., confusing speakers or confusing actions with speech about actions), or 3 = fabrication (i.e., providing details about interactions or conversations that never occurred).

Disclosures. In addition, children's disclosures of breakage were coded across interview recall question type (i.e., interaction recall questions or conversation recall questions).

Reliability. Coders individually coded the same 20% of the sample and achieved reliability on all variables. They then coded independently, coding separate interviews. For categorical variables (i.e., content and accuracy), all kappas were $>.80$. When assessing the number of details per response, intraclass correlations (ICCs) were all $>.92$.

Results

We conducted four series of analyses. First, we analyzed the amount of information that children provided in response to the interaction recall questions and conversation recall questions in three ways: overall, by content, and by conversation question. Second, we assessed accuracy in the same three manners by examining proportion accuracy scores. Third, we examined the kinds of errors made by children. Finally, we examined the extent to which children disclosed breakage across the two interviews, with special attention to whether conversation questions elicited disclosures for the first time in the second interview. Several of the variables were positively skewed. For each applicable analysis, we ran parallel analyses, first conducting a logarithmic transformation. In these instances, results were similar, and as such, the analyses on untransformed data are presented.

Analyses on amount of information. To determine whether there were effects of questioning, we conducted a series of parallel repeated-measures analyses of variance (ANOVAs). In all analyses, age was entered as a between-subjects factor, comparing the responses of younger (56% of sample; 5- to 7-year-olds) to older children (44% of sample; 8- to 9-year-olds). Question characteristics were entered as within-subjects repeated measures. The dependent measure was the number of details provided. Preliminary analyses revealed no effects of toy condition (break, no break) or order (interaction questions first, conversation questions first), and therefore those factors were excluded.

Overall recall. Descriptives are presented in Table 1. To determine whether age or questioning phase influenced the amount of recall provided, a repeated-measures ANOVA, with age entered as a between-subjects factor, was conducted on the number of details children provided to interaction recall questions compared with conversation recall questions. There was a main effect of age, $F(1, 81) = 18.06, p < .001, \eta_p^2 = .18$. Older children provided more recall ($M = 200.39, SD = 85.24$) than did younger children

($M = 132.62, SD = 60.00$). There was also a main effect for recall type, $F(1, 81) = 30.03, p < .001, \eta_p^2 = .27$. Children provided more details to interaction recall questions than to conversation recall questions (Table 1).

Action, description, and conversation details. We next conducted three parallel analyses separately analyzing the different types of recall children provided: actions, descriptions and conversations (Table 2). For action details, there were significant main effects of age, $F(1, 81) = 6.12, p = .015, \eta_p^2 = .07$, as well as recall type, $F(1, 81) = 309.95, p < .001, \eta_p^2 = .79$, and a significant Age \times Recall type interaction, $F(1, 81) = 6.21, p = .015, \eta_p^2 = .07$. In response to interaction recall questions, older children provided more action details ($M = 90.33, SD = 43.32$) than younger children ($M = 68.87, SD = 31.33$), $t(81) = 2.62, p = .011$, 95% confidence interval (CI) [5.16, 37.77]. In response to conversation recall questions, older ($M = 8.17, SD = 13.85$) and younger children ($M = 7.09, SD = 11.55$) provided similar numbers of action details, $t(81) = 0.39, p = .699$, 95% CI [-4.47, 6.63]. Both older and younger children produced more action details to interaction recall questions than conversation questions; older children $t(35) = 11.27, p < .001$, 95% CI [67.35, 96.97]; younger children $t(46) = 13.80, p < .001$, 95% CI [52.77, 70.80].

For description details, there was only a main effect of recall type, $F(1, 81) = 7.21, p < .001, \eta_p^2 = .49$; children provided more description details to interaction recall questions than to conversation recall questions (Table 2).

For conversation details, there were main effects for age, $F(1, 81) = 29.05, p < .001, \eta_p^2 = .26$, recall type, $F(1, 81) = 284.98, p < .001, \eta_p^2 = .78$, and a significant Age \times Recall type interaction, $F(1, 81) = 36.68, p < .001, \eta_p^2 = .31$. In response to interaction recall questions, older and younger children provided similar numbers of conversation details; older $M = 4.58, SD = 11.59$, younger $M = 3.96, SD = 8.80$, $t(81) = 0.28, p = .780$, 95% CI [-3.82, 5.08]. In response to conversation recall questions, older children provided twice as many conversation details ($M = 71.67, SD = 34.20$) as did younger children ($M = 35.62, SD = 19.38$), $t(81) = 6.07, p < .001$, 95% CI [24.24, 47.86]. For both older and younger children, conversation recall questions produced larger numbers of conversation details than interaction questions; older children $t(35) = 11.27, p < .001$, 95% CI [67.35, 96.97]; younger children $t(46) = 13.80, p < .001$, 95% CI [52.77, 70.80].

Conversation details by conversation question. We then focused on the number of conversation details children provided in response to each of the four conversation recall questions (Table 3). A repeated measures ANOVA was conducted on the number of conversation details as the dependent measure, with the four conversation questions entered as a within-subjects repeated measure

Table 1

Number of Details and Percentage True/False/Subjective/Indeterminate, as Well as Accuracy, by Type of Questions (Interaction Recall or Conversation Recall) in Study 1

Variable	True			False			Subjective			Indeterminate			Overall			Accuracy T/(T + F) %
	M	(SD)	%	M	(SD)	%	M	(SD)	%	M	(SD)	%	M	(SD)	%	
Interaction	74.63	(41.28)	80	5.08	(8.38)	6	4.13	(5.73)	4	10.25	(14.51)	10	94.09	(47.56)	59	93
Conversation	55.78	(39.31)	80	4.41	(6.64)	8	4.48	(6.96)	7	3.24	(7.89)	5	67.92	(41.83)	41	91
Overall	130.41	(72.03)	80	9.49	(10.84)	6	8.61	(9.67)	6	13.49	(17.05)	8	162.01	(79.13)	100	92

Table 2

Number of Details and Percentage True/False/Subjective/Indeterminate, as Well as Accuracy, by Type of Detail Content and Phase of Recall (Interaction Recall or Conversation Recall) in Study 1

Variable	True			False			Subjective			Indeterminate			Overall			Accuracy T/ (T + F) %
	M	(SD)	%	M	(SD)	%	M	(SD)	%	M	(SD)	%	M	(SD)	%	
Interaction recall phase																
Actions	62.48	(35.59)	82	4.26	(7.26)	5	2.91	(3.91)	4	8.52	(13.38)	9	78.18	(38.29)	89	93
Descriptions	5.90	(8.74)	77	.25	(1.40)	2	.68	(2.29)	13	.52	(1.80)	8	7.35	(9.46)	7	97
Conversations	3.30	(8.71)	79	.39	(1.79)	11	.28	(1.64)	7	.26	(1.81)	3	4.23	(10.04)	4	88
Conversation recall phase																
Actions	6.81	(11.97)	88	.06	(.55)	1	.00	(.00)	0	.69	(2.00)	10	7.55	(12.53)	11	98
Descriptions	3.73	(6.78)	94	.00	(.00)	0	.00	(.00)	0	.36	(2.60)	6	4.10	(7.52)	6	100
Conversations	30.23	(30.23)	76	4.47	(6.73)	9	4.46	(6.88)	10	2.09	(5.72)	5	51.25	(32.14)	83	89

(you say question, interviewer says question, stranger says question, stranger says you say question).

There were significant main effects for age, $F(1, 81) = 36.67$, $p < .001$, $\eta_p^2 = .31$, question, $F(1, 81) = 48.50$, $p < .001$, $\eta_p^2 = .38$, and a significant Age \times Question interaction, $F(3, 79) = 42.52$, $p < .001$, $\eta_p^2 = .62$. While older ($M = 5.00$, $SD = 5.58$) and younger children ($M = 4.43$, $SD = 5.23$) provided equivalent conversation details to the overt coaching question, $t(81) = 0.48$, $p = .631$, 95% CI [-1.80, 2.95], older children provided more conversation details to the question about what was said by the child ($M_{\text{older}} = 21.05$, $SD_{\text{older}} = 15.14$; $M_{\text{younger}} = 10.95$, $SD_{\text{younger}} = 10.32$), $t(81) = 3.61$, $p = .001$, 95% CI [4.53, 15.66], the interviewer ($M_{\text{older}} = 24.67$, $SD_{\text{older}} = 19.13$; $M_{\text{younger}} = 8.60$, $SD_{\text{younger}} = 6.79$), $t(81) = 5.35$, $p < .001$, 95% CI [10.09, 22.05], and the stranger ($M_{\text{older}} = 21.06$, $SD_{\text{older}} = 11.97$; $M_{\text{younger}} = 12.09$, $SD_{\text{younger}} = 8.79$), $t(81) = 3.94$, $p < .001$, 95% CI [4.44, 13.50]).

Accuracy. Parallel ANOVA models were conducted. However, the dependent measure was the proportion of accurate responses. Accuracy was calculated as the number of true details divided by the total number of details that could be classified as true or false (thus excluding subjective details and indeterminate details from the analysis; Brown & Pipe, 2003a, 2003b). To adjust for multiple comparisons, Bonferroni corrections were used. Preliminary analyses revealed no effects of toy condition (break, no break) or order (interaction questions first, conversation questions first), and therefore those factors were excluded.

Overall accuracy of recall. Table 1 reports the accuracy of recall provided by children during the interaction recall questions and conversation recall questions during the T2 interview. To determine whether age or questioning phase influenced accuracy,

a repeated-measures ANOVA was conducted on the proportion of accurate responses children provided to interaction recall questions compared with conversation recall questions. There were no significant main effects or interactions. Accuracy rates were high: interaction recall ($M = .93$, $SD = .11$), conversation recall ($M = .91$, $SD = .13$), overall, ($M = .92$, $SD = .09$).

Accuracy regarding conversation, action, and description details. Descriptive statistics are presented in Table 2. Because many children did not provide new details across all three types, we analyzed each type separately; for interaction recall questions, we examined the accuracies for action information, because that was the type of information most often elicited by interaction recall questions. Similarly, for conversation recall questions, we examined conversation accuracies. No significant differences of age emerged. Examination of the means revealed that accuracy proportion scores for both action information given during interaction recall ($M = .93$, $SD = .11$) and conversation information provided to conversation recall ($M = .89$, $SD = .16$) questions were quite high and did not vary by age. For toy description, which occurred to a comparable extent across interaction and conversation questions, 97% of the information children gave was accurate.

Accuracy of conversation details by conversation question. We then examined the accuracy of conversation details children provided in response to each of the four questions in conversation recall. Descriptive statistics are presented in Table 3. Because some children did not provide new details across all four conversation questions, we had limited power to conduct an omnibus assessment. Therefore, we focused on the coaching question compared with the other conversation questions, and conducted a paired-samples t test among the children who had accuracy proportions for both variables ($n = 58$). Children had higher accuracy

Table 3

Number of Conversation Details and Percentage True/False/Subjective/Indeterminate, as Well as Accuracy, by Conversation Question in Study 1

Variable	True			False			Subjective			Indeterminate			Overall			Accuracy T/ (T + F) %
	M	(SD)	%	M	(SD)	%	M	(SD)	%	M	(SD)	%	M	(SD)	%	
You say	12.05	(13.09)	79	1.42	(4.18)	8	.90	(2.62)	7	.96	(3.87)	6	15.34	(13.52)	28	90
Interviewer says	13.31	(15.26)	85	.39	(2.00)	3	1.45	(4.11)	9	.42	(1.96)	3	15.57	(15.69)	28	97
Stranger says	12.18	(10.11)	76	.96	(3.06)	6	2.12	(4.18)	13	.71	(2.82)	4	15.98	(11.16)	34	93
Stranger says you should say	2.47	(3.92)	53	2.11	(4.58)	45	.00	(.00)	0	.10	(.88)	2	4.67	(5.36)	10	61

scores to noncoaching conversation questions ($M = .96$, $SD = .11$) than the final conversation question about coaching ($M = .61$, $SD = .49$), $t(57) = 5.16$, $p < .001$, 95% CI [.22, .49]). Surprisingly, accuracies in response to the question about coaching were virtually identical in the no-break condition (60%) and the break condition (62%). Examining the pattern of responses more closely, we found that unlike the other conversation questions, a substantial percentage of children gave "don't know" answers to the question about coaching, and these were more common in the no-break condition (38%) than in the break condition (16%), $\chi^2(1, 83) = 4.79$, $p = .029$.

Errors made by children. To assess whether conversation recall questions increased source monitoring errors, we next examined the types of errors children made. We classified errors into three groups: (a) secrecy, (b) source confusion, or (c) fabrications. We examined errors summed across the three content areas (actions, descriptions, and conversations), and specifically examined errors for conversation details to conversation questions. Descriptives are presented in Table 4.

We conducted a series of independent-samples t tests to assess whether age (dichotomous: younger and older children) or toy break condition affected whether children would err. We looked at errors generally and at each specific type of error (secrecy, source, fabrication). There were no effects for age. Overall, children in the toy break condition made more errors, $t(81) = 2.07$, $p = .041$, 95% CI [.01, .63], and when providing conversation details, $t(81) = 2.38$, $p = .020$, 95% CI [.03, .30]. Secrecy errors were more likely in the toy break condition, $t(81) = 2.99$, $p = .004$, 95% CI [.06, .31], and when providing conversation details, $t(81) = 2.52$, $p = .014$, 95% CI [.03, .25]. Toy break condition did not affect the number of source errors or fabrications.

To understand the small number of source monitoring errors that occurred, we looked more closely at how children described their prior disclosures. We found that children often moved to general descriptions of toys during the second interview, whereby 69% of children reported general play with "toys" instead of disclosing specific toys, limiting the potential for source misattribution. Further, to understand the secrecy errors that children made, we examined how often these errors occurred in response to the overt question about coaching. Such errors occurred in response to the overt coaching question 88% of the time.

Consistency in disclosure. As we noted in the introduction, questions about conversations with suspects might lead children to

disclose transgressions for the first time. We examined the pattern of disclosure across the two interviews. For children in the *toy break* condition ($n = 43$), 27 children (63%) disclosed breakage at some point during the first interview. All of these children redisclosed breakage in the second interview. For children who were in the *no break* condition ($n = 40$), no child made a false report of breakage at any point during the first or second interview. Among the 16 children (37%) who failed to disclose breakage during the first interview, 19% of these children ($n = 3$) disclosed in response to the conversation recall questions (two of these children were 7 years old; one was 8 years old). All three children disclosed to the overt question about coaching ("What did the man say you should say to the lady?").

Discussion

The purpose of Study 1 was to assess the amount and accuracy of information that children could recall about a conversation regarding a prior event, which for half the children included a transgression. We predicted that questions explicitly asking children about their conversations would elicit conversation information missing when children were asked about their interactions with the stranger. This was supported: when children were asked interaction recall questions about the stranger, they only reported an average of three true details about their conversation, comprising less than 5% of their overall recall (Table 2). Furthermore, older children were no more productive than younger children, and because of their higher overall productivity, a smaller percentage of their recall referred to their conversation. When children were asked conversation recall questions about the stranger, they provided an average of over 12 true details about their conversations, and the other conversation recall questions reliably elicited information about children's conversation with the interviewer.

We were also interested in whether the accuracy of information children provided about their conversations would be comparable to the accuracy of information children provided about their interactions. We found that children were highly accurate to both types of recall (hovering above 90%), and there was no difference in their accuracy proportions for interaction recall compared with conversation recall. Asking interaction questions before or after conversation questions had no effect on productivity or accuracy; hence recalling the interaction did not facilitate recall of conver-

Table 4
Mean Number of Errors (for Conversation Details and Overall), and Whether Children Ever Made Such Errors, by Toy Break Condition, in Study 1

Variable	No break ($n = 43$)			Break ($n = 40$)			Overall ($n = 83$)		
	M	(SD)	% Ever erred	M	(SD)	% Ever erred	M	(SD)	% Ever erred
Conversation details									
Secrecy	.00	.00	0	.16	(.43)	14	.08	(.32)	7
Source	1.53	(3.87)	18	1.21	(3.69)	12	1.36	(3.76)	14
Fabrication	2.98	(5.99)	28	2.86	(5.85)	33	2.92	(5.88)	30
Overall details									
Secrecy	.00	.00	0	.33	(.99)	19	.17	(.73)	10
Source	3.28	(6.78)	30	2.77	(5.19)	28	3.01	(5.98)	29
Fabrication	7.48	(11.29)	50	6.07	(9.88)	44	6.75	(10.54)	47

sation questions, nor did it lead to confusion between what occurred and what was said.

When children did err, it was most frequently in response to the question explicitly inquiring into coaching (e.g., “What did the man say you should say to the lady?”). This question elicited a higher rate of error than the other questions, and also elicited the fewest number of details, probably because there was only so much to say. Among children who had not experienced a transgression (and thus had not been told to say anything to the interviewer), it elicited an unusually high percentage of “don’t know” responses, which could easily be misconstrued as evasive or reluctant.

Finally, we examined children’s disclosure patterns across the two interviews, paying special attention to whether the conversation recall questions elicited disclosures of breakage among children who had not previously disclosed. We found that nearly 20% of children who had not previously disclosed did so in response to the conversation questions, in particular, the final question overtly inquiring about coaching. However, this was the same question that elicited higher rates of inaccurate information, when compared with the other three, more general, conversation recall questions.

Study 2

There were several goals in Study 2. First, because of the small numbers, we sought to replicate the finding that conversation questions would elicit new transgression disclosures in a larger group of children who had experienced the transgression but had never disclosed. Therefore, we assessed children’s conversational memory in a sample of children who had all experienced a transgression, and we omitted yes/no questions in the first interview that explicitly asked about the transgression. Second, we explored the potential productivity and accuracy of other types of recall questions, both questions that referred more obliquely to conversations (by asking about what the child had heard) and questions that referred to conversation topics rather than conversational partners. These questions also enabled us to determine whether we could elicit transgression disclosures without overtly asking about coaching.

Some research has suggested that asking children about what they have “heard” may be effective in eliciting conversation details. Poole and Lindsay (1995) found that “heard” prompts (along with general “saw” prompts), elicited a 56% and 68% increase in correctly reported information (3- to 4-year-olds and 5- to 6-year-olds, respectively) about an interaction with a science teacher, without compromising children’s accuracy. Elischberger and Roebbers (2001) compared “heard” and “saw” questions to repeated free recall for a video. For children who did not understand the “heard” prompt, the researchers would explain that it included what people said. They found that verbal prompts about what children saw and heard elicited more information from children than general repeated recall without compromising accuracy (accuracy rates were around 98%). Unfortunately, neither study reported the amount or accuracy of conversational information specifically. Study 2 assessed the potential for the “heard” prompt to elicit conversation details.

Another potentially promising approach is to focus on the topics of conversations, as opposed to focusing entirely on speakers. In Study 1, conversation questions focused on what each speaker said

(e.g., “What did the man say? “What did the lady say?” “What did you say?”). In Study 2, we were interested in the productivity and accuracy of topical questions (e.g., “What did [the interviewer] say about the toys?” “What did [the interviewer] say about [the stranger]?”). We anticipated that the greater specificity might elicit details without increasing error, including details about the transgression.

Method

The following protocol was approved by the IRB at the University of Southern California.

Participants. The sample included seventy-one 4- to 9-year-old children ($M = 7$ years old, $SD = 1.45$, 38% male). The sample was comparable to Study 1, except with a higher proportion of Latinos (87% Latino, 9% African American, 3% Biracial).

Materials and procedure. The method for Study 2 was identical to Study 1, with five modifications. First, all of the children experienced toy breakage and were admonished to keep it a secret. Second, because of staffing changes in the research team, both the stranger and interviewer were female. This necessitated some changes in questioning to disambiguate the two adults. Third, at T1, the interviewer did not explicitly ask children about breakage. Fourth, at T2, all children received interaction questions before conversation questions. We found no order effects in Study 1, and asking interaction questions first is consistent with recommended practice for forensic interviewing.

Fifth, and most importantly, the conversation recall questions were expanded and changed. After interaction recall questions, children were presented with a picture of the stranger and asked, “Is this the lady you played with?” Then, the T2 interviewer stated, “Okay, I heard you talked to someone about what happened.” They then showed the child a picture of the previous interviewer and asked, “Is this the lady that you talked to?” All children identified both the stranger and interviewer. Children were then asked nine consecutive questions about their conversation with the stranger (five questions) and interviewer (four questions). The T2 interviewer used the picture to specify who the conversation was with, as both the stranger and interviewer were female, and pronouns would be insufficient. The order of asking about each speaker (five questions about stranger; four questions about interviewer) was counterbalanced between children. However, there were no effects observed for order of asking about one speaker before the other, and as such, this is not discussed further.

The five questions about the stranger conversation were as follows, asked in the following order. First, children were asked to “Tell me everything you heard when you played with this lady (stranger; T2 interviewer pointed to picture of the stranger).” Second, children were asked, “What did this lady (pointing to picture of stranger) say about the toys?” Third, children were asked, “What did you say to this lady (pointing to picture of stranger) about the toys?” Fourth, children were asked, “What did you say to this lady (pointing to picture of stranger) about that lady (pointing to the picture of the T1 interviewer)?” Fifth, children were asked, “What did this lady (pointing to picture of stranger) say you should say to that lady (pointing to picture of the T1 interviewer)?”

The four questions about the T1 interviewer conversation were as follows, asked in the following order. First, children were asked

to "Tell me everything you heard when you talked with that lady (pointing to picture of the T1 interviewer)." Second, children were asked, "What did that lady (pointing to the picture of the T1 interviewer) say about the toys?" Third, children were asked "What did you say to that lady (pointing to picture of T1 interviewer) about the toys?" Fourth, children were asked, "What did that lady (pointing to picture of T1 interviewer) say about this lady (pointing to picture of stranger)?"

The debriefing and coding were identical to Study 1. For categorical variables (i.e., content & accuracy), all Kappas were $>.80$. When assessing the number of details per response, ICCs were all $>.92$.

Results

Our analyses mirrored those of Study 1. As in the prior study, several of the variables were positively skewed. For each applicable analysis, we ran parallel analyses, first conducting a logarithmic transformation. In these instances, results were similar, and as such, the analyses on untransformed data are presented.

Analyses on amount of recall. To determine whether there were effects of questioning, we conducted a series of parallel repeated-measures ANOVAs. In all analyses, age was entered as a between-subjects factor, comparing the responses of younger (52% of sample; 4- to 6-year-olds) to older (48% of sample; 7- to 9-year-olds) children. Question characteristics were entered as within-subjects repeated measures. The dependent measure was the number of details provided.

Overall amount of recall. Descriptives are presented in Table 5. To determine whether age or questioning phase influenced the amount of recall provided, a repeated-measures ANOVA, with age entered as a between-subjects factor, was conducted on the number of details children provided to interaction recall questions compared with conversation recall questions. There was a main effect of age, $F(1, 69) = 23.42, p < .001, \eta_p^2 = .25$; older children provided more recall ($M = 272.65, SD = 130.60$) than did younger children ($M = 150.11, SD = 78.34$). There was also a main effect for recall type, $F(1, 69) = 24.95, p < .001, \eta_p^2 = .27$; children provided more details to interaction recall questions than to conversation recall questions (Table 5).

New action, description, and conversation details. We next conducted three parallel analyses separately analyzing the different types of recall children provided: actions, descriptions and conversations (Table 6). For action details, there were significant main effects of age, $F(1, 69) = 15.89, p < .001, \eta_p^2 = .19$, as well as recall type, $F(1, 69) = 175.89, p < .001, \eta_p^2 = .72$, and a significant Age \times Recall type interaction, $F(1, 69) = 12.97, p = .001, \eta_p^2 = .16$. In response to interaction recall questions, older

children provided more action details ($M = 122.50, SD = 61.59$) than younger children ($M = 73.14, SD = 42.16$), $t(69) = 3.97, p < .001, 95\% \text{ CI } [24.55, 74.18]$. In response to conversation recall questions, older ($M = 19.65, SD = 17.79$) and younger children ($M = 14.21, SD = 13.84$) provided a similar number of action details, $t(69) = 1.43, p = .154, 95\% \text{ CI } [-2.08, 12.94]$. Both older and younger children produced more action details to interaction recall questions than to conversation questions; older children $t(33) = 10.08, p < .001, 95\% \text{ CI } [82.10, 123.60]$, younger children $t(36) = 8.42, p < .001, 95\% \text{ CI } [44.73, 73.11]$.

For description details, there was only a main effect of age, $F(1, 69) = 8.13, p = .006, \eta_p^2 = .11$; older children provided more description details ($M = 6.47, SD = 9.55$) than did younger children ($M = 1.73, SD = 3.20$), $t(69) = 2.85, p = .006, 95\% \text{ CI } [1.42, 8.06]$.

For conversation details, there was a main effect of age, $F(1, 69) = 20.32, p < .001, \eta_p^2 = .23$, as well as recall type, $F(1, 69) = 75.49, p < .001, \eta_p^2 = .52$. Older children provided more conversation details ($M = 89.56, SD = 55.78$) than did younger children ($M = 42.14, SD = 30.13$), $t(69) = 4.51, p < .001, 95\% \text{ CI } [26.43, 68.41]$. Children provided more conversation details to conversation recall questions than they did to action recall questions, $t(70) = 8.58, p < .001, 95\% \text{ CI } [26.62, 42.74]$ (Table 6).

Conversation details by conversation question. We then focused on the number of conversation details children provided in response to each of the nine questions in conversation recall. Descriptive statistics, by question, are presented in Table 7. Here, two separate analyses were conducted: one on the five questions about the conversation with the stranger, and one on the four questions about the conversation with the previous interviewer. For each analysis, a repeated measures ANOVA was conducted on the number of new conversation details as the dependent measure, with the conversation questions entered as a within-subjects repeated measure (five stranger questions, four interviewer questions), with age entered as a between-subjects factor.

For the stranger question analyses, there was a main effect for age, $F(1, 69) = 8.86, p = .004, \eta_p^2 = .11$, as well as for question, $F(4, 66) = 4.75, p = .002, \eta_p^2 = .22$. Older children ($M = 32.35, SD = 24.23$) provided more conversation details to the conversation questions than did younger children ($M = 17.96, SD = 15.97$), $t(69) = 2.98, p = .004, 95\% \text{ CI } [4.75, 24.03]$. Pairwise comparisons, adjusted with Bonferroni's correction, revealed that children provided more details to the first question about what the child heard, than the final question about what the child was asked to say, $t(69) = 2.99, p = .032, 95\% \text{ CI } [-1.89, -9.47]$ (Table 7).

For the interviewer question analyses, there was a main effect for age, $F(1, 69) = 8.24, p < .001, \eta_p^2 = .61$, as well as for

Table 5

Number of Details and Percentage True/False/Subjective/Indeterminate, as Well as Accuracy, by Type of Questions (Interaction Recall or Conversation Recall) in Study 2

Variable	True			False			Subjective			Questionable			Overall			Accuracy T/(T + F) %
	M	(SD)	%	M	(SD)	%	M	(SD)	%	M	(SD)	%	M	(SD)	%	
Interaction	103.62	(72.29)	81	7.83	(12.52)	7	4.99	(9.67)	3	12.58	(19.35)	9	129.01	(92.67)	61	92
Conversation	58.94	(40.33)	74	2.23	(5.35)	3	6.56	(8.56)	8	12.04	(15.94)	15	79.77	(50.04)	39	96
Overall	162.56	(96.06)	78	10.06	(13.81)	5	11.55	(13.68)	5	24.62	(28.28)	12	208.79	(122.47)	100	93

Table 6
Number of Details and Percentage True/False/Subjective/Indeterminate, as Well as Accuracy, by Type of Detail Content and Phase of Recall (Interaction Recall or Conversation Recall) in Study 2

Variable	True			False			Subjective			Indeterminate			Overall			Accuracy T/ (T + F) %
	M	(SD)	%	M	(SD)	%	M	(SD)	%	M	(SD)	%	M	(SD)	%	
Interaction recall phase																
Actions	80.82	(49.78)	84	6.14	(11.29)	6	1.38	(3.59)	1	8.43	(12.88)	9	96.77	(57.61)	88	93
Descriptions	2.09	(4.14)	78	.07	(.59)	4	.36	(1.25)	10	.41	(1.86)	8	2.93	(5.38)	2	96
Conversations	13.17	(21.10)	87	.75	(2.67)	3	.10	(.83)	0	1.07	(4.79)	10	15.08	(23.09)	9	97
Conversation recall phase																
Actions	11.96	(12.67)	69	1.40	(4.87)	12	.14	(.83)	1	3.31	(6.34)	18	16.82	(15.97)	27	85
Descriptions	.72	(3.70)	60	.06	(.47)	17	.06	(.47)	17	.24	(2.02)	6	1.07	(5.55)	1	80
Conversations	41.15	(31.99)	84	.51	(1.89)	1	1.35	(3.61)	2	6.75	(12.13)	13	49.76	(36.03)	72	99

question, $F(3, 67) = 5.56, p = .002, \eta_p^2 = .20$. Older children ($M = 33.00, SD = 26.19$) provided more conversation details to the conversation questions than did younger children ($M = 18.60, SD = 14.83$), $t(70) = 2.87, p = .005, 95\% CI [4.39, 24.42]$. Pairwise comparisons, adjusted with Bonferroni's correction, revealed that children provided more details to the first question about what the child heard, than the final question about what the interviewer said about the stranger, $t(70) = 3.98, p = .001, 95\% CI [3.20, 9.63]$.

Analyses on accuracy proportions.

Overall accuracy of recall. Table 5 reports the accuracy of recall provided by children during the interaction recall questions and conversation recall questions during the T2 interview. To determine whether age or questioning phase influenced accuracy, a repeated-measures ANOVA was conducted on the proportion of accurate responses children provided to interaction recall questions compared with conversation recall questions. There were only a main effect of recall type, $F(1, 69) = 5.66, p = .02, \eta_p^2 = .08$, accuracy rates were higher to conversation recall questions than to interaction recall questions (Table 5).

Accuracy regarding conversation, action, and description details. We next conducted similar analyses separately analyzing the different types of recall children provided: actions, descriptions, or conversations. Table 6 presents descriptive information on children's accuracy by detail content. Because many children did not provide new details across all three types, we analyzed

each type separately; for interaction recall questions, we examined the accuracies for action information, since that was the type of information most often elicited by interaction recall questions. Similarly, for conversation recall questions, we examined conversation accuracies. No significant differences emerged. Examination of the means revealed that accuracy proportion scores for both action information given during interaction recall ($M = .93, SD = .12$) and conversation information provided to conversation recall ($M = .99, SD = .04$) questions were quite high and did not vary by age. For toy description, which occurred to a comparable extent across interaction and conversation questions, 95% of the information children gave was accurate ($SD = .21$).

Accuracy of conversation details by conversation question.

We then examined the accuracy of conversation details children provided in response to the questions in conversation recall (Table 7). As in Study 1, we conducted a paired-samples t test comparing the accuracy for noncoaching conversation questions to the accuracy for the final conversation question about coaching. Children had higher accuracy scores to the noncoaching conversation questions ($M = .98, SD = .10$) than the question about coaching ($M = .83, SD = .38, t(35) = 2.24, p = .031, 95\% CI [.01, .29]$). Similar to Study 1, we found that unlike the other conversation questions, a substantial percentage of children gave "don't know" answers to the question about coaching (25%).

Errors made by children. As in Study 1, we examined three kinds of errors: 1) secrecy, 2) source confusion, or 3) fabrications.

Table 7
Number of Details and Percentage True/False/Subjective/Indeterminate, as Well as Accuracy, by Conversation Questions in Study 2

Variable	True			False			Subjective			Indeterminate			Overall			Accuracy T/ (T + F) %
	M	(SD)	%	M	(SD)	%	M	(SD)	%	M	(SD)	%	M	(SD)	%	
Stranger	20.72	(19.42)	84	.43	(1.75)	2	1.20	(3.54)	5	2.50	(5.33)	10	24.86	(21.46)	50	98
Stranger heard	8.11	(14.69)	92	.00	(.00)	0	.37	(2.39)	4	.34	(2.03)	4	8.82	(15.13)	13	100
Stranger say toys	4.38	(6.25)	89	.07	(.59)	1	.20	(.97)	4	.28	(1.29)	6	4.93	(6.69)	12	97
Stranger you say	3.86	(6.29)	76	.00	(.00)	0	.55	(2.52)	11	.68	(2.53)	13	5.09	(6.66)	13	100
Stranger say interviewer (IV)	2.16	(4.41)	75	.00	(.00)	0	.00	.00	0	.73	(2.89)	25	2.89	(4.96)	6	100
Stranger say you should say	2.21	(4.88)	70	.37	(1.67)	12	.09	(.71)	3	.48	(2.41)	15	3.14	(6.32)	6	83
IV	20.58	(20.29)	81	.20	(1.13)	1	.16	(.95)	0	4.66	(10.24)	18	25.39	(22.03)	50	98
IV heard	8.79	(11.33)	88	.00	.00	0	.00	.00	0	1.54	(6.27)	12	10.09	(12.05)	22	100
IV say toys	5.48	(9.70)	78	.18	(1.13)	3	.00	.00	0	1.32	(3.52)	19	6.99	(10.10)	13	94
IV you say	3.96	(7.35)	86	.01	(.12)	0	.10	(.83)	2	.54	(2.68)	12	4.61	(7.54)	8	96
IV say Stranger	2.35	(4.71)	61	.00	(.00)	0	.06	(.48)	1	1.45	(5.38)	38	3.86	(6.70)	7	100

False conversation details were rare—descriptives are presented in Table 8. We then conducted a series of independent-sample *t* tests to assess whether age (dichotomous: younger and older children) affected whether children would provide more errors (of each kind or overall). There were no effects.

As in Study 1, in to understand the small number of source monitoring errors that occurred, we looked more closely at how children described their prior disclosures. As in Study 1, we found that children were likely to generally describe the toys, and toy play, at the second interview (68%), limiting the potential for source misattribution. Further, to understand the secrecy errors that children made, we examined how often these errors occurred in response to the overt question about coaching. Such errors occurred in response to the overt coaching question 75% of the time.

Consistency in disclosure. We examined the pattern of disclosure across the two interviews. For children who disclosed breakage at some point during the first interview ($n = 35$; 49%), all of these children redisclosed breakage in the second interview. Among the 36 children (51%) who failed to disclose breakage during the first interview, 19% of these children ($n = 7$) disclosed in response to the conversation recall questions. Hence, these children disclosed breakage for the first time when asked about their conversations. Remarkably, none of these disclosures were elicited by the question specifically asking about coaching.

Discussion

Study 2 enabled us to assess the productivity and accuracy of a different set of conversation recall questions. We started with a question about what the child “heard,” and proceeded to ask about both individual speakers and topics. There were several notable findings. First, we replicated the finding in Study 1 that asking children about conversations elicited new disclosures of the transgression. Furthermore, these disclosures occurred without resort to the explicit question about coaching. Second, we replicated the finding in Study 1 that asking children recall questions about conversations elicited conversational information that was lacking when children were asked about their interactions, and that the accuracy rates for recall questions about interactions and conversations were similarly high, except with regard to the question explicitly asking about coaching. As in Study 1, a substantial percentage of children’s errors regarded secrecy surrounding the transgression, and source monitoring errors were uncommon.

Table 8
Mean Number of Errors (for Conversation Details and Overall) in Study 2

Variable	<i>M</i>	(<i>SD</i>)	% Ever erred
Conversation details			
Secrecy	.20	(.88)	6
Source	.54	(2.10)	21
Fabrication	1.54	(4.57)	20
Overall details			
Secrecy	.20	(.88)	6
Source	2.01	(4.85)	21
Fabrication	7.96	(12.38)	49

General Discussion

We examined the potential utility of asking children recall questions about conversations. We conducted two experiments in which 4- to 9-year-old children interacted with a stranger, some experienced a transgression (with admonishments of secrecy), and all were asked to describe their interaction with the stranger in two interviews. The second interview, conducted a week after their interaction with the stranger, included conversation recall questions inquiring into children’s conversations with the stranger and with the first interviewer. Study 1 examined questions asking about what each speaker said (child, stranger, and interviewer), whereas Study 2 asked questions about what the child “heard” and questions that referenced both speaker and topic (e.g., what the stranger had said about the toys).

We predicted that explicitly asking children about their conversations would elicit unique conversation information. This prediction was supported: recall questions asking children about their interactions with the stranger produced small amounts of conversational information, whereas recall questions asking about conversations reliably elicited conversational recall. Children’s conversational recall was highly accurate, and comparable to their recall of interactions. However, questions explicitly asking about coaching were subject to higher rates of error, leading many children to fabricate details, falsely deny coaching, and give “don’t know” responses. In both studies, older children recalled more than younger children, but accuracies did not vary by age.

Finally, we examined children’s disclosure of the transgression across the two interviews, and although the numbers were small, we consistently found that about 20% of children disclosed the transgression for the first time when asked about their conversations. Although Study 1 suggested that these additional disclosures could only be elicited with a specific reference to coaching, Study 2 found that less specific questions about conversations with the stranger were similarly productive. Conversation questions may play a role in encouraging children to disclose wrongdoing that they are otherwise reluctant to reveal. Notably, however, a substantial percentage of children had failed to disclose the transgression at the end of the procedure, highlighting the importance of finding nonsuggestive means of overcoming children’s reluctance to disclose.

With respect to types of errors, we thought that children, and younger children in particular, might have difficulty with source monitoring. Younger children might be prone to confuse what they had disclosed to the first interviewer with what they had experienced and/or discussed with the stranger. Source monitoring errors are most common when the sources are highly similar (Lindsay, Johnson, & Kwon, 1991). Here, children interacted with the stranger and the first interviewer in the same room and in quick succession. Conducting an immediate interview likely facilitated children’s memory for the interaction with the stranger, but by the same token it increased the similarity of the interview and the interaction. Nevertheless, source monitoring errors were quite rare.

Limitations and Future Directions

Several factors would likely lead to increases in source monitoring errors. In addition to the usual factors (such as delay), we believe three are particularly important for conversational memory. First, by only asking recall questions, we may have minimized

the likelihood of source confusion. Children were inclined to describe their conversations with the first interviewer in general terms, such as referring to disclosing play with “the toys” rather than specific toys. Much of the work finding developmental differences in source monitoring asks recognition questions (e.g., Foley et al., 1983). Had we asked children yes/no questions whether they had disclosed play or breakage with specific toys, error rates probably would have increased (Lawson & London, 2017), and of course overtly suggestive questions (e.g., “You told the lady the alligator broke, didn’t you?”) would probably further exacerbate error. Yes/no questions about conversations are common when children are questioned in court (Stolzenberg & Lyon, 2014).

Second, because children were thoroughly questioned about their interaction with the stranger in the first interview, there was less opportunity to subsequently confuse what had been disclosed with what had only been experienced. Children’s first disclosures of abuse are likely to have been to family members or friends rather than forensic interviewers, and thus be incomplete. This creates a potential for children to falsely remember disclosing details in their initial disclosures when subsequently questioned.

Third, children were only required to distinguish between two conversations, one they had with the stranger and another they had with the first interviewer. In actual abuse cases, children are routinely interviewed on multiple occasions, and in abuse trials, children are often asked to distinguish among multiple interviews (Stolzenberg & Lyon, 2014). Children are likely to develop script memories of the content of repeated interviews, and have difficulty in distinguishing among them, much like they have difficulties in recalling a single instantiation of a repeated event (Roberts & Powell, 2001). These difficulties highlight the importance of asking children conversation questions as early as possible in the investigative process.

There is little doubt that children’s accuracy would have suffered had we insisted on verbatim recall. As noted in the introduction, a classic finding is that verbatim recall for speech disappears almost instantaneously (Sachs, 1967). Neither practice guides recommending conversation questions nor legal rules regarding the admissibility of hearsay require that statements be recalled verbatim. However, in child sexual abuse trials, a recurrent issue is the suggestiveness of adult’s questions (Stolzenberg & Lyon, 2014), and question-form is unlikely to be retained in gist memory. Both laboratory and observational research has demonstrated the difficulty adults have in recalling the form of their questions (Bruck, Ceci, & Francoeur, 1999; Korkman, Jussola, & Santtila, 2014; Lamb, Orbach, Sternberg, Hershkowitz, & Horowitz, 2000; Warren & Woodall, 1999), and children are no better (Lawson & London, 2015, 2017). Because even contemporaneous verbatim note-taking by interviewers fails to capture the types of questions asked or the quantity of details provided (Lamb et al., 2000), practitioners are advised to record interviews whenever possible.

Because children’s conversational memory is an understudied topic, there is a considerable amount of work that needs to be done. Our finding that conversational recall questions helped to elicit details missing from children’s recall of their interactions is consistent with a broader literature examining ways in which specific topics important in child forensic interviews can be fruitfully elicited through focused recall questions, such as children’s emo-

tional reactions to abuse (Lyon, Scurich, Choi, Handmaker, & Blank, 2012). However, even when children make only minimal reference to conversations when asked about interactions, as they did in our studies, persistent use of cued invitations (e.g., “Tell me more about what he said”) might be productive, and should be explored in future work, along with systematic examination of the comparative productivity and accuracy of different types of conversation questions, including general references (“What did you hear?” and “What did people say?”), references to specific speakers, and references to specific topics. Moreover, counterbalancing of question order would allow better estimation of the incremental utility of specific questions.

Problems associated with the question we used to specifically refer to coaching (“What did [the stranger] say you should say to [the interviewer]?”) are worthy of future work in order to understand why they elicited higher error rates overall and higher rates of “don’t know” responses among children who had not experienced a transgression. We suspect that the high rate of “don’t know” responses reflects the suggestiveness of the question, which presumes that the stranger told the child to say something, but doesn’t suggest what that something is. Although a “don’t know” response is less worrisome than suggested or fabricated information, it could be misinterpreted as resistance by a child reluctant to disclose coaching.

The coaching question may also have been difficult because of its syntactic form. From a linguistic perspective, the question “What did the man say you should say?” involves second-order embedding. Statements reporting speech involve first-order embedding; for example, “The man said playing was fun” contains a single embedded clause “playing was fun” (O’Grady, 1997). Statements reporting coached speech involve second-order embedding; for example, “The man said I should say playing was fun” embeds “playing was fun” within “I should say” which in turn is embedded within “the man said.” Research primarily interested in children’s understanding of mental states (which has explored a parallel distinction between first-order and second-order theory of mind tasks) has found that whereas young preschool children have difficulty with first-order embedded statements about speech (Perner, Sprung, Zauner, & Haider, 2003), young grade schoolchildren continue to struggle understanding second-order embedded statements (Hollebrandse, Hobbs, De Villiers, & Roeper, 2008).

Because of the importance of assessing adult influence, future work should consider other ways in which linguistic and cognitive immaturity may adversely affect children’s reports. For example, children are routinely asked about asking and telling in child sexual abuse trials (Stolzenberg & Lyon, 2014), and they must distinguish between requests and commands (whether suspects asked or told them to perform various acts) and between questions and answers (whether disclosure recipients asked or told them information). The problem is that children appear to initially understand telling as synonymous with saying, and view asking as a type of telling (Stolzenberg, McWilliams, & Lyon, 2017a). Hence, they will endorse that an adult told them “what happened” when the adult only asked them “What happened?” Questions about adult influences may also be referentially ambiguous. For example, children are sometimes asked whether an adult has told them “what to say.” A child who was simply advised to “tell the truth” might answer affirmatively, creating the false impression that he or she was coached (Lyon, 2013).

Future work should also assess the possible effects of children's socioeconomic status, ethnicity, and native language on their conversational memory. Our sample was comprised of predominantly low-income ethnic minority children, many of whom were likely from Spanish-speaking homes. As such, they are quite unlike the middle-class Caucasian children who predominate in much developmental research. However, given the changing demographics of the United States, they are quite representative; a majority of births are racial and ethnic minorities, and a majority of children of all ages are projected to be racial and ethnic minorities by 2020 (Lichter & Qian, 2018). Similarly, almost half (44%) of children under nine years of age in the United States are low income (Koball & Jiang, 2018).

In sum, the results provide tentative support for including recall questions about conversations in forensic interviews with children. These questions may elicit useful information about children's interactions with suspects and with disclosure recipients and may encourage children to disclose information they are motivated to conceal. At the same time, questions directly inquiring into possible coaching appear problematic, and future work is needed to identify the optimal means for eliciting complete reports from children about their conversations.

References

- Ahern, E. C., & Lamb, M. E. (2017). Children's reports of disclosure recipient reactions in forensic interviews: Comparing the NICHD and MoGP. *Journal of Police and Criminal Psychology, 32*, 85–93. <http://dx.doi.org/10.1007/s11896-016-9205-x>
- Ahern, E. C., & Lyon, T. D. (2011). *Supplemental investigative interview questions*. Los Angeles, CA: Author.
- American Professional Society on the Abuse of Children. (2012). *Forensic interviewing in cases of suspected abuse*. Chicago, IL: Author.
- Bowen, C. J., & Howie, P. M. (2002). Context and cue cards in young children's testimony: A comparison of brief narrative elaboration and context reinstatement. *Journal of Applied Psychology, 87*, 1077–1085. <http://dx.doi.org/10.1037/0021-9010.87.6.1077>
- Brown, D., & Pipe, M. E. (2003a). Individual differences in children's event memory reports and the narrative elaboration technique. *Journal of Applied Psychology, 88*, 195–206. <http://dx.doi.org/10.1037/0021-9010.88.2.195>
- Brown, D., & Pipe, M. E. (2003b). Variations on a technique: Enhancing children's recall using narrative elaboration training. *Applied Cognitive Psychology, 17*, 377–399. <http://dx.doi.org/10.1002/acp.876>
- Brubacher, S. P., Roberts, K. P., & Powell, M. (2012). Retrieval of episodic versus generic information: Does the order of recall affect the amount and accuracy of details reported by children about repeated events? *Developmental Psychology, 48*, 111–122. <http://dx.doi.org/10.1037/a0025864>
- Bruck, M., Ceci, S. J., & Francoeur, E. (1999). The accuracy of mothers' memories of conversations with their preschool children. *Journal of Experimental Psychology: Applied, 5*, 89–106. <http://dx.doi.org/10.1037/1076-898X.5.1.89>
- Camparo, L. B., Wagner, J. T., & Saywitz, K. J. (2001). Interviewing children about real and fictitious events: Revisiting the narrative elaboration procedure. *Law and Human Behavior, 25*, 63–80. <http://dx.doi.org/10.1023/A:1005691926064>
- Cassel, W. S., Roebbers, C. E., & Bjorklund, D. F. (1996). Developmental patterns of eyewitness responses to repeated and increasingly suggestive questions. *Journal of Experimental Child Psychology, 61*, 116–133. <http://dx.doi.org/10.1006/jecp.1996.0008>
- Connolly, D. A., & Gordon, H. M. (2014). Can order of general and specific memory prompts help children to recall an instance of a repeated event that was different from the others? *Psychology, Crime & Law, 20*, 852–864. <http://dx.doi.org/10.1080/1068316X.2014.885969>
- Davis, D., & Friedman, R. D. (2007). Memory for conversations: The orphan child of witness memory research. In M. P. Toglia, J. D. Read, D. F. Ross, & R. C. L. Lindsay (Eds.), *Handbook of eyewitness psychology: Vol. 2. Memory for events* (pp. 3–57). Mahwah, NJ: Erlbaum.
- Davis, D., Kemmelmeier, M., & Follette, W. C. (2005). Conversational memory on trial. In Y. I. Noy & W. Karwowski (Eds.), *Handbook of human factors in litigation* (pp. 12–1–29–12). Boca Raton, FL: CRC Press.
- Dorado, J. S., & Saywitz, K. J. (2001). Interviewing preschoolers from low- and middle-SES communities: A test of the narrative elaboration recall improvement technique. *Journal of Clinical Child Psychology, 30*, 568–580. http://dx.doi.org/10.1207/S15374424JCCP3004_12
- Duke, S. B., Lee, A. S., & Pager, C. K. W. (2007). A picture's worth a thousand words: Conversational versus eyewitness testimony in criminal convictions. *The American Criminal Law Review, 44*, 1–52. Retrieved from http://digitalcommons.law.yale.edu/fss_papers/829
- Elishberger, H. B., & Roebbers, C. M. (2001). Improving young children's free narratives about an observed event: The effects of nonspecific verbal prompts. *International Journal of Behavioral Development, 25*, 160–166. <http://dx.doi.org/10.1080/01650250042000203>
- Foley, M. A. (2014). Children's memory for source. In P. J. Bauer & R. Fivush (Eds.), *The Wiley handbook on the development of children's memory, Volume III* (pp. 427–452). Chichester, United Kingdom: Wiley.
- Foley, M. A., Johnson, M. K., & Raye, C. L. (1983). Age-related changes in confusion between memories for thoughts and memories for speech. *Child Development, 54*, 51–60. <http://dx.doi.org/10.2307/1129860>
- Goodman, G. S., Quas, J. A., Bulkley, J., & Shapiro, C. (1999). Innovations for child witnesses: A national survey. *Psychology, Public Policy, and Law, 5*, 255–281. <http://dx.doi.org/10.1037/1076-8971.5.2.255>
- Greenstock, J., & Pipe, M. E. (1996). Interviewing children about past events: The influence of peer support and misleading questions. *Child Abuse & Neglect: The International Journal, 20*, 69–80. [http://dx.doi.org/10.1016/0145-2134\(95\)00117-4](http://dx.doi.org/10.1016/0145-2134(95)00117-4)
- Hershkowitz, I., & Elul, A. (1999). The effects of investigative utterances on Israeli children's reports of physical abuse. *Applied Developmental Science, 3*, 28–33. http://dx.doi.org/10.1207/s1532480xads0301_4
- Hershkowitz, I., Lamb, M. E., Orbach, Y., Katz, C., & Horowitz, D. (2012). The development of communicative and narrative skills among preschoolers: Lessons from forensic interviews about child abuse. *Child Development, 83*, 611–622.
- Hollebrandse, B., Hobbs, K., De Villiers, J. L., & Roeper, T. (2008). Second order embedding and second order false belief. In A. Gavarro & M. J. Freitas (Eds.), *Language acquisition and development, Proceedings of GALA 2007* (pp. 270–280). Newcastle, United Kingdom: Cambridge Scholars Publishing.
- Home Office. (1992). *Memorandum of good practice on video-recorded interviews with child witnesses for criminal proceedings*. London, United Kingdom: Her Majesty's Stationery Office.
- Koball, H., & Jiang, Y. (2018). *Basic facts about low-income children: Children under 9 years, 2016*. New York, NY: National Center for Children in Poverty.
- Korkman, J., Juusola, A., & Santtila, P. (2014). Who made the disclosure? Recorded discussions between children and caretakers suspecting child abuse. *Psychology, Crime & Law, 20*, 994–1004. <http://dx.doi.org/10.1080/1068316X.2014.902455>
- Kulkofsky, S. (2010). The effects of verbal labels and vocabulary skill on memory and suggestibility. *Journal of Applied Developmental Psychology, 31*, 460–466. <http://dx.doi.org/10.1016/j.appdev.2010.09.002>
- Lamb, M. E., Hershkowitz, I., Orbach, Y., & Esplin, P. W. (2008). *Tell me what happened: Structured investigative interviews of child victims*

- and witnesses. London, England: Wiley. <http://dx.doi.org/10.1002/9780470773291>
- Lamb, M. E., Orbach, Y., Sternberg, K. J., Hershkowitz, I., & Horowitz, D. (2000). Accuracy of investigators' verbatim notes of their forensic interviews with alleged child abuse victims. *Law and Human Behavior, 24*, 699–708. <http://dx.doi.org/10.1023/A:1005556404636>
- Lawson, M., & London, K. (2015). Tell me everything you discussed: Children's memory for dyadic conversations after a 1-week or a 3-week delay. *Behavioral Sciences & the Law, 33*, 429–445. <http://dx.doi.org/10.1002/bsl.2184>
- Lawson, M., & London, K. (2017). Children's memory for conversations after a 1-year delay. *Journal of Applied Research in Memory & Cognition, 6*, 328–336. <http://dx.doi.org/10.1016/j.jarmac.2017.07.001>
- Leclerc, B., Proulx, J., & Beaugregard, E. (2009). Examining the modus operandi of sexual offenders against children and its practical implications. *Aggression and Violent Behavior, 14*, 5–12. <http://dx.doi.org/10.1016/j.avb.2008.08.001>
- Lichter, D. T., & Qian, Z. (2018). Children at risk: Diversity, inequality, and the third demographic transition. In D. Poston, Jr. (Ed.), *Low fertility regimes and demographic and societal change*. Cham, Switzerland: Springer. http://dx.doi.org/10.1007/978-3-319-64061-7_10
- Lindsay, D. S., Johnson, M. K., & Kwon, P. (1991). Developmental changes in memory source monitoring. *Journal of Experimental Child Psychology, 52*, 297–318. [http://dx.doi.org/10.1016/0022-0965\(91\)90065-Z](http://dx.doi.org/10.1016/0022-0965(91)90065-Z)
- Lyon, T. D. (2013). Child witnesses and imagination: Lying, hypothetical reasoning, and referential ambiguity. In M. Taylor (Ed.), *The Oxford handbook of the development of imagination* (pp. 126–136). New York, NY: Oxford.
- Lyon, T. D., Scurich, N., Choi, K., Handmaker, S., & Blank, K. (2012). "How did you feel?": Increasing child sexual abuse witnesses' production of evaluative information. *Law & Human Behavior, 36*, 448–457. <http://dx.doi.org/10.1037/h0093986>
- Lyon, T. D., & Stolzenberg, S. N. (2014). Children's memory for conversations about sexual abuse: Legal and psychological implications. *Roger Williams Law Review, 19*, 411–450.
- Malloy, L. C., Brubacher, S. P., & Lamb, M. E. (2013). "Because she's one who listens": Children discuss disclosure recipients in forensic interviews. *Child Maltreatment, 18*, 245–251. <http://dx.doi.org/10.1177/1077559513497250>
- Marsh, E. J. (2007). Retelling is not the same as recalling: Implications for memory. *Current Directions in Psychological Science, 16*, 16–20. <http://dx.doi.org/10.1111/j.1467-8721.2007.00467.x>
- Mueller, C. B., & Kirkpatrick, L. C. (2015). *Evidence under the rules*. New York, NY: Wolters Kluwer Law & Business.
- Myers, J. E. B. (2016). *Myers on evidence of interpersonal violence: Child maltreatment, intimate partner violence, rape, stalking, and elder abuse* (6th ed.). New York, NY: Wolters & Kluwer.
- Nathanson, R., Crank, J. N., Saywitz, K. J., & Ruegg, E. (2007). Enhancing the oral narratives of children with learning disabilities. *Reading & Writing Quarterly: Overcoming Learning Difficulties, 23*, 315–331. <http://dx.doi.org/10.1080/10573560701501610>
- O'Grady, W. (1997). *Syntactic development*. Chicago, IL: University of Chicago. <http://dx.doi.org/10.7208/chicago/9780226620787.001.0001>
- Perner, J., Sprung, M., Zauner, P., & Haider, H. (2003). Want that is understood well before say that, think that, and false belief: A test of de Villiers's linguistic determinism on German-speaking children. *Child Development, 74*, 179–188. <http://dx.doi.org/10.1111/1467-8624.t01-1-00529>
- Peskin, J., & Ardino, V. (2003). Representing the mental world in children's social behavior: Playing hide and seek and keeping a secret. *Social Development, 12*, 496–512. <http://dx.doi.org/10.1111/1467-9507.00245>
- Peterson, C., Warren, K. L., & Hayes, A. H. (2013). Revisiting narrative elaboration training with an ecologically relevant event. *Journal of Cognition and Development, 14*, 154–174. <http://dx.doi.org/10.1080/15248372.2011.638688>
- Poole, D. A., & Lindsay, D. S. (1995). Interviewing preschoolers: Effects of nonsuggestive techniques, parental coaching, and leading questions on reports of nonexperienced events. *Journal of Experimental Child Psychology, 60*, 129–154. <http://dx.doi.org/10.1006/jecp.1995.1035>
- Principe, G. F., & Schindewolf, E. (2012). Natural conversations as a source of false memories in children: Implications for the testimony of young witnesses. *Developmental Review, 32*, 205–223. <http://dx.doi.org/10.1016/j.dr.2012.06.003>
- Revised Investigative Interview Protocol. (2014). Retrieved from <http://nichdprotocol.com/wp-content/uploads/2013/03/RevisedProtocolTMWH2final-1.pdf>
- Roberts, K. P., & Powell, M. B. (2001). Describing individual incidents of sexual abuse: A review of research on the effects of multiple sources of information on children's reports. *Child Abuse & Neglect, 25*, 1643–1659. [http://dx.doi.org/10.1016/S0145-2134\(01\)00290-3](http://dx.doi.org/10.1016/S0145-2134(01)00290-3)
- Sachs, J. S. (1967). Recognition memory for syntactic and semantic aspects of connected discourse. *Attention, Perception & Psychophysics, 2*, 437–442. <http://dx.doi.org/10.3758/BF03208784>
- Saywitz, K. J., & Snyder, L. (1996). Narrative elaboration: Test of a new procedure for interviewing children. *Journal of Consulting and Clinical Psychology, 64*, 1347–1357. <http://dx.doi.org/10.1037/0022-006X.64.6.1347>
- Saywitz, K. J., Snyder, L., & Lamphear, V. (1996). Helping children tell what happened: A follow-up study of the narrative elaboration procedure. *Child Maltreatment, 1*, 200–212. <http://dx.doi.org/10.1177/1077559596001003003>
- Smith, B. E., & Elstein, S. G. (1993). *The prosecution of child sexual abuse cases: Final report*. Washington, DC: National Center for Child Abuse & Neglect.
- Sodian, B., & Wimmer, H. (1987). Children's understanding of inference as a source of knowledge. *Child Development, 58*, 424–433. <http://dx.doi.org/10.2307/1130519>
- Stafford, L., & Daly, J. A. (1984). Conversational memory: The effects of time, recall, mode, and memory expectancies on remembrances of natural conversations. *Human Communication Research, 14*, 203–229. <http://dx.doi.org/10.1111/j.1468-2958.1987.tb00127.x>
- Stolzenberg, S. N., & Lyon, T. D. (2014). How attorneys question children about the dynamics of sexual abuse and disclosure in criminal trials. *Psychology, Public Policy, and Law, 20*, 19–30. <http://dx.doi.org/10.1037/a0035000>
- Stolzenberg, S. N., McWilliams, K., & Lyon, T. D. (2017a). Ask versus tell: Potential confusion when child witnesses are questioned about conversations. *Journal of Experimental Psychology: Applied, 23*, 447–459. <http://dx.doi.org/10.1037/xap0000136>
- Stolzenberg, S. N., McWilliams, K., & Lyon, T. D. (2017b). The effect of the hypothetical putative confession and negatively valenced yes/no questions on maltreated and non-maltreated children's disclosures of a minor transgression. *Child Maltreatment, 22*, 167–173. <http://dx.doi.org/10.1177/1077559516673734>
- Warren, A. R., & Woodall, C. E. (1999). The reliability of hearsay testimony: How well do interviewers recall their interviews with children? *Psychology, Public Policy, and Law, 5*, 355–371. <http://dx.doi.org/10.1037/1076-8971.5.2.355>

Received April 2, 2018

Revision received May 11, 2018

Accepted May 13, 2018 ■