Children's Memory for Conversations After a 1-Year Delay

Monica Lawson*
University of Notre Dame, United States

Kamala London
University of Toledo, United States

Forensic professionals frequently ask children to recount the occurrence and the content of allegation-relevant conversations during maltreatment investigations. However, little is known about children's conversation memory, especially after long delays that often characterize forensic investigations. Participants included 77 9-year-olds. When children were 8 years old, they participated in two to-be-remembered conversations: a target conversation and an initial interview regarding the target conversation. Memory for both conversations was examined 1 year later. After a year, children remembered the topic of the target conversation, but gist recall of statements was limited. Additionally, children demonstrated a yes-bias when answering yes/no questions about conversation statements. Virtually none of the children recalled participating in the initial interview. Our results suggest that after 1 year, children may remember the topic of seminal conversations, but memory for conversational statements may be sparse and unreliable. Furthermore, children may not recall engaging in peripheral conversations (such as interviews) after extended delays.

General Audience Summary

Children involved in maltreatment investigations discuss abuse with a variety of people in both formal and informal settings. Forensic interviewers and attorneys typically question children about prior conversations related to the maltreatment allegations. However, little is known about children's memory for conversations in which they actively participated, especially after long delays that often characterize forensic investigations. Seventy-seven children participated in two to-be-remembered conversations: a target conversation with a storyteller and an initial interview for the target conversation. Children's memory for both conversations was examined one year later during a follow-up interview. Children were eight years old when the conversations occurred, and nine years old during the 1-year follow-up interview. Following the 1-year delay, children accurately remembered the general topic of the target conversation, but they rarely remembered specific statements that were uttered. Virtually no children remembered talking to the initial interviewer. Additionally, when asked whether certain statements were said during the target conversation, children were inclined to agree that statements were said, even when they were not said during the target conversation. Our results suggest that after one year, children may remember the topic of seminal conversations, but memory for conversational statements may be sparse and unreliable. Furthermore, children may not recall engaging in peripheral conversations (such as interviews) after extended delays.

Author Note

* Correspondence concerning this article should be addressed to Monica Lawson, University of Notre Dame, 118 Haggar Hall, Notre Dame, IN 46556, United States. Contact: mlawson5@nd.edu

In almost any child maltreatment investigation, children's conversations with their alleged perpetrators and with their disclosure recipients are evaluated. Because conversations can exert a powerful influence on children's event reports (Principe & Schindewolf, 2012) and on children's maltreatment allegations (Schaeffer, Leventhal, & Asnes, 2011), examining the content of allegation-relevant conversations is essential to understanding abuse dynamics, for testing alternative hypotheses, and for conducting reliability assessments of children's testimony (Lyon & Stolzenberg, 2014; National Children's Advocacy Center, 2016). Forensic interviewers and attorneys routinely ask children to recount allegation-relevant conversations and often inquire about specific statements that may have been said (Ahern & Lamb, 2016; Hershowitz, Lanes, & Lamb, 2007; Malloy, Brubacher, & Lamb, 2013; Stolzenberg & Lyon, 2014).

However, there is a dearth of research examining the reliability of children's conversation memory, especially after long delays. Maltreatment investigations often are lengthy and delayed disclosure of abuse is common (London, Bruck, Ceci, & Shuman, 2005). Allegation-relevant conversations often occur months or years before formal investigations where children are asked to recount these interactions. Although children often reliably report details from highly stressful experiences (Peterson, 2012) and staged events (London, Bruck, & Melynk, 2009; Price & Connolly, 2013) months after they occur, whether children accurately recount conversations from the distant past is unknown. The primary impetus of the current study was to examine: (a) children's recollection for the occurrence of conversations, and (b) children's recall and recognition memory for conversational statements after a 1-year delay.

Davis and Friedman (2007) called memory for conversation the "orphan child of eyewitness memory researchers" to emphasize the neglect of research in this area. A handful of studies have examined adults' memory for conversations with another adult (Stafford, Burggraf, & Sharkey, 1987; Stafford & Daly, 1984) or with a child (Bruck, Ceci, & Francoeur, 1999; Lamb, Orbach, Sternberg, Hershkowitz, & Horowitz, 2000; Warren & Woodall, 1999). Although children's memory for individual sentences and short stories has been examined (e.g., Greenhoot, Beyer, & Curtis, 2014; Reyna & Kiernan, 1994), children's memory for conversations is arguably distinct from children's memory for passively perceived verbal stimuli and is deserving of its own investigation (Davis & Friedman, 2007; Davis, Kemmelmeier, & Follette, 2005; Duke, Lee, & Pager, 2007).

Conversations are cognitively demanding in requiring that participants attend to and interpret often fragmented statements declared by their conversational partner, in combination with tone and body language, while simultaneously preparing responses that adhere to conversational maxims (Grice, 1975). The dyadic nature of conversations requires encoding, storing, and retrieving not only statements, but also the declarant of each remark. Additionally, the repetitiveness of daily discourse might make sourcing memories for specific statements difficult. Consequently, memory for conversations might be especially vulnerable to erroneous encoding, rapid decay, and source monitoring errors (Davis & Friedman, 2007; Davis et al., 2005; Duke et al., 2007). Moreover, children might not understand interviewer and attorney prompts inquiring about conversations (Evans, Stolzenberg, Lee, & Lyon, 2014; Stolzenberg et al., 2017).

Empirically supported methods for asking children about conversations do not currently exist although testing hypotheses regarding third-party influence and asking about informal disclosures are ubiquitous interviewing practices. Forensic interviewers are universally advised to use open-ended questions before yes/no questions because children provide more accurate and more elaborative responses to open-ended versus yes/no prompts (Lamb, Orbach, Hershowitz, Esplin, & Horowitz, 2007). Children often recollect some statements from dyadic conversations when prompted with open-ended questions (Stolzenberg & Lyon, 2015), but recall is limited in accuracy and in completeness after relatively short delays. In our previous study (Lawson & London, 2015), eight-year-olds accurately recounted 7% and 4% of a dyadic conversation after a 1- or 3-week delay. Accuracy was judged by gist rather than for verbatim content. Children's limited conversation recall, in conjunction with evidence suggesting that many child witnesses do not spontaneously recount allegation-relevant conversations in forensic contexts (Ahern & Lamb, 2016; Malloy et al., 2013; Stolzenberg & Lyon, 2014), highlights potential challenges in eliciting information about conversations with open-ended questions.

Many questions posed by forensic professionals about conversations can be answered with a "yes" or a "no" response (Ahern & Lamb, 2016; Stolzenberg & Lyon, 2014). However, relying on yes/no questions for eliciting conversation information might be problematic. Children usually answer yes/no questions even when they lack the necessary information to respond and are advised that "I don't know" can be an appropriate answer (Fritzley & Lee, 2003; Waterman & Blades, 2013). Children are more reluctant to admit a lack of knowledge after extended delays than during more immediate questioning (Waterman & Blades, 2013). Furthermore, children might demonstrate a response bias in answering yes/no questions, especially when questions are confusing, the subject of the question is unfamiliar, and/or after experiencing a delay (Fritzley & Lee, 2003; Fritzley, Lindsay, & Lee, 2013). Asking false yes/no questions (i.e., questions where the correct answer is no) might be particularly problematic because children often acquiesce to false questions even when the content concerns potentially stressful experiences such as a dentist appointment (Rocha, Marche, & Briere, 2013) or physical trauma (Peterson & Biggs, 1997). Children are more inclined to agree with false questions months after the to-be-remembered event occurs than soon after the experience (London et al., 2009; Rocha et al., 2013). In our prior study, children who experienced a 1-week delay from the target conversation to the initial interview answered 74% of the yes/no questions correctly. Children who experienced a 3-week delay were significantly less accurate in answering yes/no questions (68% of questions were answered correctly) compared to children who experienced a 1-week delay. Whether children demonstrated a response bias was not examined.

The current investigation expands upon our original study by reevaluating children's conversation memory after a 1-year

delay. Children participated in two conversations in the original study: (a) a target conversation and (b) an initial interview. We anticipated that, after a 1-year delay, children would recall engaging in the target conversation, but not the initial interview. Informal observations of forensic interviews and trial testimony for expert witness consultation provided by the second author (KL) demonstrates that children often deny having had earlier conversations about the allegations, despite evidence to the contrary. Children's recognition memory for specific statements and whether children demonstrate response biases in answering yes/no questions after a 1-year delay were of particular interest in the current report.

Method

Participants

Ninety eight-year-olds ($M = 8.07$ years, $SD = 0.31$) participated in our original study. Approximately one year after participating in our initial study ($M = 329$ days, $SD = 86$), 80 of the 90 original participants were relocated, and parental consent was obtained for participation in the present study. Children were nine years old ($M = 8.91$ years, $SD = 0.40$) at the time of the 1-year follow-up interview that is of interest in the current report. The 10 children we were unable to contact did not significantly differ from the 80 children included in the follow-up sample in gender, age, or measures of conversation memory during the initial interview (F's < 2.48, p's $> .12$). Three additional children were excluded from the current analysis. A fire drill interrupted follow-up interviewing for one child. Two children provided extremely off-topic follow-up reports. The final sample for the current study consisted of 77 nine-year-olds (53% male; 92% White). Eight-year-olds were invited to participate in our original study because children this age are typically fluent conversationalists who are able to verbally contribute to conversations and provide narratives (Fivush & Haden, 1997). Additionally, children of varying ages were not included in the original sample to bolster linguistic similarity across target conversations.

Procedure

Target conversation. As part of our original study, children individually participated in a target conversation with a female researcher (i.e., the storyteller). The target conversation was conceptualized as a "storytelling event" in which children verbally created a story with the storyteller about a monkey playing baseball. The goal of the storytelling event was to engage children in a novel conversation with an adult to minimize the possibilitiy that their later memory reports were based on pre-existing schemas for routine conversations. Additionally, the current paradigm was deliberately designed to exclude a physical staged event that is typical in child eyewitness research to allow for an examination of children's conversation memory as independently as possible from children's memory for physical aspects of events. Prior to engaging in the target conversation, all children listened to a two-minute recorded story about a racecar driving dinosaur.

The storytellers were highly trained to deliver 25 scripted questions (e.g., "How old is your monkey?") during the target conversations in order to guide the flow of the dialogue and to increase similarity across all target conversations. Children responded to the storytellers' questions in any manner they saw fit. During the target conversation, the storyteller wore a tall yellow top hat to aid children in identifying the target conversation during subsequent interviews. The average target conversation lasted approximately five minutes.

Initial interview following a 1- or 3-week delay. Children's memory for the target conversation was assessed by an unfamiliar researcher after a delay of 1- or 3-weeks (with random assignment). After building rapport, children were instructed, with open-ended prompts, to recount the target conversation (i.e., "Tell me your story exactly the same way you told the storyteller"). Additional facilitators were used to encourage exhaustive reports (e.g., "Tell me more"). Then, children were asked whether the storyteller asked questions to help them create their story. Children were prompted to recall any questions that the storyteller asked (e.g., "Tell me one question she asked you") as well as their original response for each question they recalled (e.g., "What did you tell her?") until children indicated that they did not remember any more questions or that their report was exhaustive.

Then, children were asked 30 yes/no questions (15 true and 15 false) about specific statements that may or may not have been said during the target conversation. Ten of the questions (5 true, 5 false) concerned statements that the child may have said during the conversation. True yes/no questions regarding children's statements were tailored for each participant according to statements that they declared during the target conversation following a standardized template. The five false yes/no questions regarding children's statements were standard for all participants. Children were also asked 10 questions (5 true, 5 false) regarding statements that the storyteller may have said. All yes/no questions regarding the storytellers' statements were standard for all participants. Additionally, children were asked 10 questions (5 true, 5 false) about question–answer pairs that may have occurred during the target conversation. Question–answer pair items captured a dyadic exchange between the storyteller and the child and consisted of a question asked by the storyteller and the child's response. True question–answer pair items were verbatim reiterations of question–answer exchanges from the target conversation. For false question–answer pair items, the verbatim content of the storytellers' question was altered to incorrectly depict the original question. These questions were originally designed to evaluate children's memory for the form of adults' questions, which is often of interest to forensic professionals due to concerns of parental coaching and use of suggestive questions (Bruck et al., 1999). However, because children performed poorly on these questions during the initial interview (Lawson & London, 2015) and because forensic professionals often ask about individual statements that may have been said by specific individuals rather than question–answer pairs (Stolzenberg & Lyon, 2014), children's performance on

question–answer pair items was not examined in the current report.

The order of questions within each of the three yes/no question categories (i.e., child statements, storyteller statements, and question–answer pairs) was random, and the presentation of the three categories was counterbalanced across participants. Forty-two children who participated in the current study experienced an initial 1-week delay between the target conversation and the initial interview, and 35 children experienced an initial delay of three weeks.

Follow-up interview following a 1-year delay. Children participated in a second interview with an unfamiliar researcher approximately one year after the target conversation. After building rapport, interviewers asked children if they talked to the storyteller. Children were reminded that the storyteller wore a tall yellow hat and was from the nearby university. The follow-up interview was identical to the initial interview with the addition of asking children whether they have ever talked to anybody else about the target conversation after children answered the yes/no questions. If children recounted previously discussing the target conversation with somebody else, they were asked with whom they discussed the target conversation and to report what was said during that conversation. These questions were designed to explore whether children had any recollection for the initial interview and mirrored the style of questioning often used by forensic professionals.

Coding

Target conversations. Audio recordings of the target conversations, the initial interviews, and the follow-up interviews were transcribed verbatim and checked for accuracy by a second researcher. All statements declared during the target conversations were segmented into utterances. Utterances were defined as a statement containing a verb and a pause (see Bruck et al., 1999). For example, the statement, "The monkey hit the ball and ran" was coded as two utterances (i.e., [1] The monkey hit the ball [2] and ran). The declarant of each utterance was coded. On average, target conversations consisted of 101 utterances ($SD = 28$), of which 50 ($SD = 23$) were declared by the child.

Interviews. Children's free recall reports were parcelled into utterances. Utterances were coded for accuracy by referencing transcripts of the target conversations. Correct utterances matched at least the gist of an original utterance stated during the participants' target conversation. For example, the utterance, "the monkey played a lot" was a correct recollection of the original utterance, "the monkey played every day." Verbatim recall was not necessary for accuracy. Utterances were classified as incorrect if they did not contain the gist of an original utterance in the target conversation. Two researchers coded utterances for the initial and follow-up interviews. Agreement between the researchers was excellent as measured by two-way single measure intraclass correlations ($ICC > .98$). The number of "I don't know," "yes," and "no" responses were tallied for true and false yes/no questions.

Results

Recall for the Occurrence and the Content of Conversations

During the 1-year follow-up interview, the majority of children (90%) correctly recalled that they talked to the storyteller about a monkey playing baseball, indicating that the target conversation was salient to most children. However, when children were asked whether they ever talked to anyone else about the target conversation, only one child indicated a previous discussion with the initial interviewer. Nine children reported that they previously told a family member or a friend about the target conversation. Because only one child recalled participating in the initial interview, children's memory for the content of the initial interview was not examined further.

In the following analyses, we examined children's recall for utterances from the target conversation. To examine whether reports differed across the initial and the follow-up interviews, and to account for the potential influence of the timing of the initial interview (see Pipe, Sutherland, Webster, Jones, & La Rooy, 2004), a series of 2 (Initial Delay: 1-week, 3-week) by 2 (Interview: initial, follow-up) mixed ANOVAs with repeated measures on the last factor were conducted on (a) the number of correct utterances recalled, (b) the number of incorrect utterances recalled, and (c) the percent of utterances from the target conversation that children correctly recounted. Several of the variables were positively skewed. However, analyses with logarithmic transformed data derived similar results as analyses using untransformed data. Furthermore, the parametric assumption of homogeneity of variance was not violated when models were conducted with untransformed data. Therefore, analyses were conducted with untransformed data.

The initial delay by interview interaction on correct utterances was significant, $F(1, 75) = 8.67$, $p < .01$, $\eta_p^2 = .10$. A Bonferroni corrected t-test (with alpha set at .0125) revealed a marginally significant difference between children who experienced the initial 1-week and 3-week delay on the number of correct utterances recounted during the initial interview, $t(75) = 2.50$, $p = .015$, $d = 0.58$. During the follow-up interview, the number of correct utterances recounted did not significantly differ by initial delay condition, $t(75) = -.009$, $p = .99$. Children in both delay conditions recounted significantly more correct utterances during the initial interview compared to the follow-up interview, $t(41) = 8.89$, $p < .001$, $d = 0.53$ for children who experienced an initial 1-week delay, and $t(34) = 5.97$, $p < .001$, $d = 0.55$ for children in the 3-week delay condition. The results revealed a main effect of interview on incorrect utterances, $F(1, 75) = 4.58$, $p = .04$, $\eta_p^2 = .06$. Children recounted significantly more incorrect utterances during the initial interview compared to the follow-up interview. Additionally, there was a main effect of interview on the percent of utterances from the target conversation that children correctly recalled, $F(1,75) = 109.35$, $p < .001$, $\eta_p^2 = .59$. Children correctly recounted more of the target conversation during the initial interview compared to the follow-up interview. See Table 1. Because recall memory was limited during the initial and the follow-up interviews, and because legal personnel rely heavily on yes/no questions when asking children about conversations, the remaining focus

Table 1
Mean (SD) Number of Utterances Provided During the Initial and Follow-Up Interviews by Initial Delay Condition

	Initial interview			Follow-up interview		
	Initial 1-week delay ($n = 42$)	Initial 3-week delay ($n = 35$)	Overall ($N = 77$)	Initial 1-week delay ($n = 42$)	Initial 3-week delay ($n = 35$)	Overall ($N = 77$)
Correct utterances†	8.14 (4.98)ab	5.57 (3.84)ac	6.97 (4.65)	2.31 (2.12)b	2.31 (2.37)c	2.31 (2.23)
Incorrect utterances	4.71 (5.90)	3.66 (3.90)	4.23 (5.09)a	3.38 (3.29)	2.71 (2.81)	3.07 (3.08)a
% of conversation	7.09 (4.41)	6.32 (4.62)	7.18 (4.54)a	2.18 (2.00)	2.36 (2.40)	2.27 (2.18)a

Note. Superscripts a, b, c indicate significant differences between categories within the same row.
†Significant interview by initial delay interaction. The difference in the number of correct utterances recalled during the initial interview by children in the 1-week and 3-week delay condition was marginally significant when alpha was adjusted for familywise error.

of our results concerns children's performance on yes/no questions.

Recognition Memory for Conversation Statements

Two children (both in the initial 1-week delay condition) were asked the wrong yes/no questions during the follow-up interview and were excluded from the following analysis. Occasionally, children were not asked some true yes/no questions regarding their own statements because the storyteller did not deliver lines during the target conversation that were used to create the true yes/no questions regarding children's statements following the standardized template. A total of 12 true yes/no questions about statements said by the child were not asked across the entire sample for this reason. Children's performance on yes/no questions was operationalized as the percent of questions that children answered correctly for true and false questions separately (i.e., the percent of "yes" responses for true questions and the percent of "no" responses for false questions). For children who were not asked all five of the true yes/no questions about statements they said during the target conversation, the denominator used to compute the percent of true questions correctly answered was adjusted to reflect the number of true questions posed. Children's performance on true and false yes/no questions concerning child versus storyteller statements did not significantly differ (p's > .12). Therefore, performance on child and storyteller yes/no questions was collapsed.

Before presenting results regarding the percent of yes/no questions children accurately answered, the results regarding "I don't know" responses are provided. In total, 22 children said "I don't know" to at least one yes/no question during the initial and/or follow-up interview. Sixty-three "I don't know" responses were provided in total, with the majority (65%) occurring during the follow-up interview. However, children rarely said "I don't know" to yes/no questions, comprising 4% of all responses. A 2 (Interview: initial, follow-up) by 2 (Question Veracity: true, false) McNemar test was conducted to determine whether the frequency of "I don't know" responses differed by true and false yes/no questions or across interviews. The McNemar test revealed the number of "I don't know" responses provided by children was not dependent on the veracity of yes/no questions nor the timing of the interview ($p = .72$). Furthermore, the number of "I don't know" responses provided by children

was not associated with children's performance on true and false yes/no questions (p's > .09). In the following analyses, "I don't know" responses were coded as inaccurate.

To examine the percent of yes/no questions that children correctly answered, a 2 (Initial Delay: 1-week, 3-weeks) by 2 (Interview: initial, follow-up) by 2 (Question Veracity: true, false) ANOVA with repeated measures on the last two factors was conducted. The results revealed two significant two-way interactions. A significant interaction between interview and initial delay condition emerged, $F(1, 73) = 8.17$, $p < .01$, $\eta_p^2 = .10$. A Bonferroni adjusted post hoc t-test with performance collapsed by question veracity revealed that during the initial interview, children who experienced a 1-week delay correctly answered significantly more yes/no questions compared to children who experienced a 3-week delay, $t(73) = 2.64$, $p = .01$, $d = 0.61$. However, performance during the follow-up interview did not significantly differ by initial delay condition, $t(62.99) = -0.68$, $p = .50$. Furthermore, children correctly answered significantly more yes/no questions during the initial interview compared to the follow-up interview, $t(39) = 10.34$, $p < .001$, $d = 1.64$ for children who experienced an initial 1-week delay, $t(34) = 3.56$, $p = .001$, $d = 0.61$ for children who experienced an initial 3-week delay.

A significant interaction between question veracity and interview also emerged, $F(1, 73) = 15.71$, $p < .001$, $\eta_p^2 = .18$. During the initial interview, performance did not significantly vary by question veracity, $t(74) = 0.14$, $p = .89$. However, during the follow-up interview, children correctly answered significantly more true questions than false questions, $t(74) = 4.85$, $p < .001$, $d = 0.56$. Additionally, performance on true yes/no questions was stable across the initial and the follow-up interviews, $t(74) = 1.77$, $p = .08$. However, children correctly answered significantly fewer false yes/no questions during the follow-up interview compared to the initial interview, $t(74) = 7.42$, $p < .001$, $d = 0.86$ (see Table 2).

Our results indicate that children correctly answered significantly more yes/no questions during the initial interview compared to the follow-up interview. During the initial interview, children who experienced a 1-week delay correctly answered significantly more yes/no questions compared to children in the 3-week delay condition, but performance did not vary by question veracity for children in either delay condition. Additionally, the initial delay experienced by children did

Table 2

Mean (SD) Percent Yes/No Questions Correctly Answered by Interview, Initial Delay Condition, and Question Veracity

	Initial interview			Follow-up interview		
	Initial 1-week delay ($n = 40$)	Initial 3-week delay ($n = 35$)	Overall ($N = 75$)	Initial 1-week delay ($n = 40$)	Initial 3-week delay ($n = 35$)	Overall ($N = 75$)
True items	80.64 (17.56)	79.95 (14.54)	80.32 (16.12)	74.31 (14.34)	78.04 (19.07)	76.05 (16.70)[*]
False items	85.75 (14.48)	73.14 (22.46)	79.87 (19.56)[a]	57.25 (22.30)	57.14 (24.80)	57.20 (23.34)[a*]
Overall performance[†]	83.19 (11.17)[ab]	76.55 (10.53)[ac]	80.09 (11.31)	65.78 (9.82)[b]	67.59 (12.92)[c]	66.62 (11.33)

Note. Superscripts a, b, c indicate significant differences between categories within the same row.
[*]Significant differences between rows. [†]Significant interview by initial delay interaction.

not influence performance during the follow-up interview. The percent of true questions that children accurately answered did not significantly vary across interviews, but children answered significantly fewer false questions correctly after the extended delay compared to the initial interview.

To examine whether children demonstrated a response bias in answering yes/no questions, a percentage score representing the number of questions to which children responded "yes" by the total number of yes/no questions children were asked was computed for the initial and follow-up interviews. One-sample t-tests were conducted to compare the percent of "yes" responses children provided during the initial and follow-up interviews to 50%, the percentage of yes/no questions children would reply "yes" to if they did not demonstrate a response bias. Mean values significantly higher than 50% indicate a significant yes-bias, whereas mean values significantly lower than 50% indicate a significant nay-saying bias (i.e., a tendency to say "no" and/or "I don't know" to yes/no questions). Children did not display a significant response bias during the initial interview ($M = 49.95\%$, $SD = 14.10\%$), $t(74) = -0.03$, $p = .98$. However, children demonstrated a significant yes-bias during the follow-up interview ($M = 59.31\%$, $SD = 16.92\%$), $t(74) = 4.77$, $p < .001$. Additionally, the initial delay condition did not significantly influence the percent of yes/no questions children responded "yes" to during the initial interview ($M = 47.17\%$, $SD = 11.74\%$ for children in the 1-week delay condition, $M = 53.13\%$, $SD = 15.81\%$ for children in the 3-week delay condition), $t(62.11) = -1.83$, $p = .07$. Children who experienced an initial 1-week ($M = 58.39\%$, $SD = 16.05\%$) or 3-week delay ($M = 60.35\%$, $SD = 18.04\%$) did not significantly differ in the percent of yes/no questions they responded "yes" to during the follow-up interview, $t(73) = -0.50$, $p = .62$.

Discussion

Consistent with our expectations, after the 1-year delay, children recalled participating in the target conversation, but not the initial interview. Additionally, children were inclined to agree with yes/no questions about conversation statements. Each finding and its forensic implications are further discussed below.

Recall for the Occurrence and the Content of Conversations

Children remembered the basic premise of the target conversation, but had difficulty recalling statements from the dialogue one year later. Although verbatim recollection was not required for accuracy, floor level recall for conversation statements was not surprising because gist recall was limited during the initial interview (as reported in Lawson & London, 2015). Additionally, few children reported telling others about the target conversation. We designed the target conversation to be memorable and unique. However, our data suggest that the discussion about the target conversation (the initial interview) was not memorable to children.

Recognition Memory for Conversation Statements

Consistent with previous research (e.g., Waterman & Blades, 2013), children rarely stated they did not know the answer to yes/no questions, suggesting that children are willing to answer yes/no questions about conversation statements. Additionally, after a 1-year delay, children had a tendency to respond "yes" to yes/no questions regardless of question veracity. This yes-bias was reflected by stability in true question performance across interviews and increased acquiesce to false questions during the follow-up interview. False yes/no questions contained inaccurate statements, but were consistent with the general theme of the discussion. After one year passed and verbatim recollection for the conversational content likely faded, children may have been more likely to agree with false questions because the question content was consistent with the general gist of the target conversation, which the majority of children recalled (see also Rocha et al., 2013). Additionally, previous exposure to false questions during the initial interview might have increased children's familiarity with these items, and children may have misattributed such statements to the target conversation versus the initial interview (see Brainerd & Reyna, 1996; London et al., 2009). Future research is needed to examine differential effects of question familiarity and response biases on children's yes/no performance when children are exposed to repeated questioning and experience extended delays. We were unable to evaluate this distinction because we did not include new yes/no questions during the follow-up interview.

Limitations

This study generated several novel findings with important implications for forensic investigations involving child witnesses, but direct application of our findings to the forensic arena should be cautioned. As with any analogue study, compromises between ecological and internal validity were made. The emotional salience of the to-be-remembered event, or conversations in our case, likely does not compare to emotions children experience in forensic contexts. However, our design allowed for carefully controlled conversations and interviews, which is crucial for establishing a foundation for research in this novel area. Future research is necessary to differentiate children's memory for emotional versus more mundane conversations. Additionally, our research was conducted among nine-year-olds, and we did not assess developmental trends. Evaluating whether age influences children's conversation memory is imperative before providing definitive recommendations for forensic professionals who talk to children of varying ages.

Furthermore, because our studies are among the first to examine children's memory for conversations and empirical guidelines for questioning children about conversations do not exist, it is important to note that children's reports may be more or less reliable when questions are formatted differently (Evans et al., 2014; Stolzenberg et al., 2017). Given that prior conversations are almost always a central issue in child maltreatment investigations, developing empirically supported questions about conversations should be a top research priority. Understanding how memory for conversation compares to memory for other stimuli will be important for creating empirically based methods for questioning children about conversations.

Practical Applications

Forensic professionals need to be aware of the strengths and the weaknesses of children's conversation memory. Although children often report having previously told one or more people about allegations during forensic interviews (Ahern & Lamb, 2016; Hershowitz et al., 2007; Malloy et al., 2013), our results suggest that children's forensic reports might underestimate children's conversation history. Children might remember more seminal allegation-relevant conversations after long delays, but memory for peripheral conversations (e.g., suggestive interviews) might be minimal. Additionally, recall for conversation statements should be expected to be limited in gist accuracy and in completeness. Furthermore, because children often do not understand that memory is suggestible (London, Bruck, Poole, & Melnyk, 2011), it is unlikely that young children would encode and later recall the leading nature of suggestive questions.

Our findings also suggest that yes/no questions about conversation statements should be carefully used, especially when the conversation may have occurred in the distant past. In the current study, children had a tendency to say "yes" to yes/no questions about conversation statements regardless of question veracity after a 1-year delay. In forensic contexts, the content of yes/no questions is often based on speculation because corroborative evidence of informal conversations rarely exists. Therefore, asking children yes/no questions about specific statements might serve to confirm preexisting beliefs rather than to uncover reliable details. Risks associated with inaccurate or incomplete conversation evidence are tremendous in situations where children engage in highly suggestive, coercive informal conversations and inaccurately recount these interactions to investigators. Children's forensic reports might appear spontaneous and reliable when in fact children's testimony was tainted by social influence during earlier conversations.

Our results indicate that eliciting reliable details about conversations from children might be challenging. Given that information about allegation-relevant conversations is imperative to many investigations involving child witnesses, future research is gravely needed to inform forensic professionals of how to best prompt children to recount conversations. We hope, together, these studies spawn further scientific inquiry into this orphan of eyewitness research.

Conflict of Interest Statement

The authors declare no conflict of interest.

Author Contributions

The design of the experiment was a joint collaboration between Monica Lawson and Kamala London. Monica Lawson was responsible for performing the experiment and analyzing and interpreting the data. The writing of the manuscript was a joint effort between both researchers.

Acknowledgements

References

Ahern, B., & Lamb, M. E. (2016). Children's reports of disclosure recipient reactions in forensic interviews: Comparing the NICHD and MoGP protocols. *Journal of Police and Criminal Psychology*, 1–9. http://dx.doi.org/10.1007/s11896-016-9205-x

Brainerd, C. J., & Reyna, V. F. (1996). Mere memory testing creates false memories in children. *Development Psychology*, *32*, 467–476. http://dx.doi.org/10.1037/0012-1649.32.3.467

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

Davis, D., & Friedman, R. D. (2007). Memory for conversation: The orphan child of witness memory researchers. In M. P. Toglia, J. D. Read, D. F. Ross, & R. C. L. Lindsay (Eds.), *Handbook of eyewitness psychology: Memory for events* (pp. 3–52). Mahwah, NJ: Lawrence Erlbaum Associates.

Davis, D., Kemmelmeier, M., & Follette, W. C. (2005). Memory for conversation on trial. In Y. I. Noy, & W. Karwowski (Eds.), *Handbook of human factors in litigation* (pp. 1–29). Boca Raton, FL: CRC Press.

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. *American Criminal Law Review*, *44*, 1–52. Retrieved from http://www.law.georgetown.edu/academics/law-journals/aclr/

Evans, A., Stolzenberg, S. N., Lee, K., & Lyon, T. D. (2014). Young children's difficulty with indirect speech acts: Implications for questioning child witnesses. *Behavioral Sciences & the Law*, *32*, 768–775. http://dx.doi.org/10.1002/bsl.2142

Fivush, R., & Haden, C. A. (1997). Narrating and representing experiences: Preschoolers' developing autobiographical accounts. In P. van den Broek, P. J. Bauer, & T. Bourg (Eds.), *Developmental spans in event representation and comprehension: Bridging functional and actual events* (pp. 169–198). Mahwah, NJ: Lawrence Erlbaum Associates.

Fritzley, V. H., & Lee, K. (2003). Do young children always say yes to yes-questions? A metadevelopmental study of the affirmation bias. *Child Development*, *74*, 1297–1313. http://dx.doi.org/10.1111/1467-8624.00608

Fritzley, V. H., Lindsay, R. C. L., & Lee, K. (2013). Young children's response tendencies toward yes-no questions concerning actions. *Child Development*, *84*, 711–725. http://dx.doi.org/10.1111/cdev.12006

Greenhoot, A. F., Beyer, A. M., & Curtis, J. (2014). More than pretty pictures? How illustrations affect parent–child story reading and children's story recall. *Frontiers in Psychology*, *5*, 1–10. http://dx.doi.org/10.3389/fpsyg.2014.00738

Grice, H. P. (1975). Logic and conversations. In P. Cole, & J. L. Morgan (Eds.), *Speech acts* (Vol. 3) *Syntax and semantics* (pp. 41–58). New York: Academic Press.

Hershowitz, I., Lanes, O., & Lamb, M. E. (2007). Exploring the disclosure of child sexual abuse with alleged victims and their parents. *Child Abuse & Neglect*, *31*, 111–123. http://dx.doi.org/10.1016/j.chiabu.2006.09.004

Lamb, M. E., Orbach, Y., Hershowitz, I., Esplin, P. W., & Horowitz, D. (2007). Structured forensic interview protocols improve the quality and informativeness of investigative interviews with children: A review of research using the NICHD Investigative Interview Protocol. *Child Abuse & Neglect*, *31*, 1201–1231. http://dx.doi.org/10.1016/j.chiabu.2007.03.021

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. Retrieved from http://www.jstor.org/stable/1394494

Lawson, M., & London, K. (2015). Tell me everything you discussed: Children's memory for a dyadic conversation after a 1-week or 3-week delay. *Behavioral Sciences and the Law*, *33*, 429–445. http://dx.doi.org/10.1002/bsl.2184

London, K., Bruck, M., Ceci, S. J., & Shuman, D. W. (2005). Disclosure of child sexual abuse: What does the research tell us about the ways that children tell? *Psychology, Public Policy, and Law*, *11*, 194–226. http://dx.doi.org/10.1037/1076-8971.11.1.194

London, K., Bruck, M., & Melynk, L. (2009). Post-event information affects children's autobiographical memory after one year. *Law and Human Behavior*, *33*, 344–355. http://dx.doi.org/10.1007/s10979-008-9147-7

London, K., Bruck, M., Poole, D., & Melnyk, L. (2011). The development of metasuggestibility in children. *Applied Cognitive Psychology*, *25*, 146–155. http://dx.doi.org/10.1002/acp.1653

Lyon, T. D., & Stolzenberg, S. N. (2014). Children's memory for conversations about sexual abuse: Legal and psychological implication. *Roger Williams University Law Review*, *19*, 411–450. Retrieved from http://rogerwilliamslawreview.org/

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

National Children's Advocacy Center. (2016). *Alternative hypotheses in the child forensic interview: What are we talking about? A research-to-practice summary*. Huntsville, AL: Author.

Peterson, C. (2012). Children's autobiographical memories across the years: Forensic implications of childhood amnesia and eyewitness memory for stressful events. *Developmental Review*, *32*, 287–306. http://dx.doi.org/10.1016/j.dr.2012.06.002

Peterson, C., & Biggs, M. (1997). Interviewing children about trauma: Problems with 'specific' questions. *Journal of Traumatic Stress*, *10*, 279–290. http://dx.doi.org/10.1002/jts.2490100208

Pipe, M. E., Sutherland, R., Webster, N., Jones, C., & La Rooy, D. (2004). Do early interviews affect children's long-term event recall? *Applied Cognitive Psychology*, *18*, 823–839. http://dx.doi.org/10.1002/acp.1053

Price, H. L., & Connolly, D. A. (2013). Suggestibility effects persist after one year in children who experienced a single or repeated event. *Journal of Applied Research in Memory and Cognition*, *2*, 89–94. http://dx.doi.org/10.1016/j.jarmac.2013.03.001

Principe, G., & 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

Reyna, V. F., & Kiernan, B. (1994). Development of gist versus verbatim memory in sentence recognition: Effects of lexical familiarity, semantic content, encoding instructions, and retention interval. *Developmental Psychology*, *30*, 178–191. http://dx.doi.org/10.1037/0012-1649.30.2.178

Rocha, E. M., Marche, T. A., & Briere, J. L. (2013). The effect of forced-choice questions in children's suggestibility: A comparison of multiple-choice and yes/no questions. *Canadian Journal of Behavioral Science*, *45*, 1–11. http://dx.doi.org/10.1037/a0028507

Schaeffer, P., Leventhal, J. M., & Asnes, A. G. (2011). Children's disclosure of sexual abuse: Learning from direct inquiry. *Child Abuse & Neglect*, *35*, 343–352. http://dx.doi.org/10.1016/j.chiabu.2011.01.014

Stafford, L., Burggraf, C. S., & Sharkey, W. F. (1987). 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

Stafford, L., & Daly, J. A. (1984). Conversational memory: The effects of recall mode and memory expectancies on remembrances of natural conversations. *Human Communication Research*, *10*, 379–402. http://dx.doi.org/10.1111/j.1468-2958.1984.tb00024.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., & Lyon, T. D. (2015). The consistency and accuracy of children's reports of an adult transgression across multiple interviews. In *American Psychology – Law Society Conference*.

Stolzenberg, S. N., McWilliams, K., & Lyon, T. D. (2017). Ask versus tell: Potential confusion when child witnesses are questioned about conversations. *Journal of Experimental Psychology: Applied*.

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

Waterman, A. H., & Blades, M. (2013). The effect of delay and individual differences on children's tendency to guess. *Developmental Psychology*, *49*, 215–226. http://dx.doi.org/10.1037/a0028354