1. Interviewing child witnesses: The effect of forced confabulation on event memory

Stacia N. Stolzenberg, University of Southern California
Kathy Pezdek, Claremont Graduate University
Interviewing child witnesses: The effect of forced confabulation on event memory

Stacia Stolzenberg *, Kathy Pezdek

Department of Psychology, Claremont Graduate University, Claremont, CA 91711, USA

ARTICLE INFO

Article history:
Received 10 January 2012
Revised 8 September 2012
Available online 23 October 2012

Keywords:
Eyewitness memory
Forced confabulation
Child witnesses
Children’s memory
Suggestibility
Developmental differences

ABSTRACT

Age differences in rates of forced confabulation and memory consequences thereof were assessed using a recall task similar to real forensic interview procedures. Children viewed a target video and were tested with the same 18 questions immediately afterward and 1 week later. Of the 18 questions, 12 were answerable; the 6 unanswerable questions referred to information not in the video. Participants in the voluntary confabulation condition had a “don’t know” response option; those in the forced confabulation condition did not. Although 6-year-olds and 9-year-olds were equally likely to provide a response to an unanswerable question initially, 1 week later 9-year-olds were significantly more likely than 6-year-olds to repeat their initial confabulated responses. These findings suggest that pressing child witnesses to answer questions they are initially reluctant to answer is not an effective practice, and the consistency of children’s responses over time is not necessarily an indication of the accuracy of their eyewitness memory.

Published by Elsevier Inc.

Introduction

Psychologists have long been familiar with the fallibility of eyewitness memory, and the flood of preschool abuse cases during the 1980s prompted eyewitness memory researchers to direct their attention to the reliability of eyewitness memory by children (Blandon-Gitlin & Pezdek, 2009). One psychological factor known to affect the accuracy of eyewitness memory for both adults and children is the extent to which suggestive influences intervene between when an eyewitness observes an event and when the eyewitness testifies about the event. Numerous studies have demonstrated that
eyewitnesses can be misled by postevent suggestions following an observed event, a finding known as the misinformation effect (Blandon-Gitlin & Pezdek, 2009; Loftus, Miller, & Burns, 1978; Pezdek, 1977).

Most of the research on memory suggestibility has examined how memory can be affected by external sources of information such as suggestive questions by interviewers (see Pezdek & Hinz, 2002, and specifically Blandon-Gitlin & Pezdek, 2009, and Ceci & Bruck, 1993, for reviews of this research with children). However, several recent studies have examined a related but different issue, namely, how eyewitnesses’ memory is affected by self-generated internal sources of information resulting from forced confabulation. Forced confabulation occurs when an individual erroneously incorporates into his or her memory for an event self-generated information that was not actually part of that event (Pezdek, 2008). Forced confabulation can occur in forensic interviews when an interviewer presses an eyewitness to answer a question even though the eyewitness has indicated that he or she does not know or is unsure of the answer to the question. Of interest in the current study was the extent to which information that was initially confabulated by a child will be remembered as part of the original event and repeated in subsequent interviews.

In a recent study on forced confabulations by Pezdek, Sperry, and Owens (2007), adults viewed a videotaped event and immediately thereafter answered 16 answerable and 6 unanswerable questions about the event (most of which were open questions). Unanswerable questions referred to information consistent with but not actually presented in the video. Participants in the voluntary confabulation condition had a “don’t know” response option; those in the forced confabulation condition did not. One week later, the same questions were answered with a “don’t know” option available for everyone. When participants were forced to confabulate answers to unanswerable questions at Time 1, 40% of the Time 2 responses in Experiment 1 and 38% of the Time 2 responses in Experiment 2 were the same answer to the question that had been generated at Time 1 even though the “don’t know” response option was available at Time 2. That is, participants misremembered events about which they had previously confabulated answers and they repeated their confabulated answers in subsequent interviews. These findings are consistent with those reported by other researchers (Hastie, Landsman, & Loftus, 1978; Zaragoza, Payment, Ackil, Drivdahl, & Beck, 2001). Furthermore, in a more recent study using a signal detection paradigm, Gombos, Pezdek, and Haymond (2012) confirmed that the forced confabulation effect is a real memory effect above and beyond the effects of response bias; forcing eyewitnesses to guess or speculate can actually change their memory.

As suggested by Pezdek et al. (2007), two cognitive stages are hypothesized during the forced confabulation process. First, the individual must retrieve at Time 2 his or her self-generated answer to the unanswerable question from Time 1. Second, the individual must make a source monitoring error at Time 2 and confuse his or her self-generated answer with information actually observed (Johnson, Hashtroudi, & Lindsay, 1993; Lindsay, 1990; Lindsay & Johnson, 1989). In numerous studies, it has been demonstrated that children are more susceptible than adults to source monitoring errors (Foley & Johnson, 1985; Foley, Johnson, & Raye, 1983; Lindsay, Johnson, & Kwon, 1991; Parker, 1995). The current study assessed whether, in addition to the effects of source monitoring, developmental differences in the persistence of forced confabulation are also affected by memory for the responses initially self-generated.

Research by Ackil and Zaragoza (1998) suggests that in a forced confabulation task, children are more likely than adults to err and misattribute to an observed target event self-generated Time 1 responses to unanswerable questions. These authors examined the postevent memories of 6-year-olds, 9-year-olds, and adults who viewed a target video and immediately afterward answered a series of questions. Each question referred to either information presented in the video (answerable questions) or information not presented in the video (unanswerable questions). Half of the participants were told to respond to every question (the forced confabulation condition); half were told to respond “don’t know” to questions they had difficulty in answering (the control condition). One week later, participants were reinterviewed and asked yes/no questions about the source of each test item (i.e., was the information from the video or the experimenter?). Ackil and Zaragoza reported that the proportion of source misattribution errors (i.e., erroneous “yes” responses to “saw it in the video” questions) was significantly higher in the forced confabulation condition than in the control condition for each age group. Furthermore, the size of this difference in misattribution errors interacted with age; source
misattribution errors for confabulated items occurred more frequently for 6-year-olds than for 9-year-olds and more frequently for 9-year-olds than for adults.

Although these results of Ackil and Zaragoza (1998) suggest that in an interview forced confabulations are more likely to be remembered and subsequently recalled by younger children than by older children, in fact several factors influence this conclusion. As suggested above, the first cognitive step during this process is that the individual must retrieve at Time 2 his or her self-generated answer to the unanswerable question from Time 1. Cued recall is one cognitive process that has been consistently reported to improve with age in children. This has been reported both in traditional experimental tasks (Bousfield, Esterson, & Whitmarsh, 1958; Cole, Frankel, & Sharp, 1971) and in eyewitness memory studies (Cohen & Harnick, 1980; Shrimpton, Oates, & Hayes, 1998). From these findings, it could be predicted that younger children would actually be less likely than older children to remember at Time 2 their self-generated confabulated responses from Time 1. To the extent that memory retrieval of previous confabulated responses plays a role in forced confabulation, younger children would be expected to exhibit lower rates of repetition of their confabulated responses than would older children and, thus, have lower rates of forced confabulation. On the other hand, as indicated above, to the extent that source monitoring errors play a role in forced confabulation, younger children, for whom source monitoring errors are typically higher, would be expected to exhibit higher rates of forced confabulation. The task used in the current study allows us to examine the combined effect of these two aspects of forced confabulation. The task used by Ackil and Zaragoza (1998)—a yes/no response regarding the source of each test item—specifically assessed age differences in source misattribution errors for information that was initially forced confabulated but did not actually assess age differences in the persistence or repeated recall of information that was initially forced confabulated.

Following the procedures of Pezdek et al. (2007), 6-year-olds and 9-year-olds viewed a target video and were interviewed immediately afterward and 1 week later. Both interviews consisted of the same 18 questions; of these, 12 were answerable questions and 6 were unanswerable questions that referred to information not in the video. In the initial interview, participants in the voluntary confabulation condition had a “don’t know” response option; those in the forced confabulation condition did not. In the second interview 1 week later, the “don’t know” response option was available for everyone. Thus, the procedure was similar to that used by Ackil and Zaragoza (1998) except that whereas those authors used only a source memory test at Time 2, in the current study both the initial interview and the delayed interview were the same recall test. The current study begins to assess the role of memory recall above and beyond the effects of source monitoring in determining the impact of forced confabulation on children. The focus of this study was on the confabulated responses to unanswerable questions, specifically (a) whether the rate of providing confabulated responses at Time 1 differs between 6-year-olds and 9-year-olds and (b) whether 1 week later 9-year-olds are more likely than 6-year-olds to repeat their Time 1 confabulated responses.

Methods

Participants

A total of 110 children participated, including 55 6-year-olds (mean age = 6.43 years, 28 male and 27 female) and 55 9-year-olds (mean age = 9.36 years, 26 male and 29 female). Children were recruited from a private elementary school in Los Angeles County in the western United States. Among this initial sample, 14 children were excluded from analyses because they had previously seen the target video and 1 child was excluded for having difficulty in answering questions. The final sample size was 95 participants (49 6-year-olds and 46 9-year-olds) for all analyses. Children who spoke English as a second language were ineligible. This was a 2 (forced or voluntary confabulation condition) × 2 (6-year-olds or 9-year-olds) between-participants factorial design. Participants were selected at these ages to mimic prior research assessing age differences in source monitoring abilities (Ackil & Zaragoza, 1998). Power analyses were conducted using methods suggested by Faul, Erdfelder, Lang, and Buchner (2007). For each analysis, the anticipated effect size was specified at .35 with alpha = .05. The sample size indicated slightly exceeds that required to detect effects with power = .90.
Materials and procedures

The procedures were adapted from Pezdek et al. (2007). Children participated in two sessions lasting approximately 15 min each. In Session 1, each participant was individually taken from the classroom and presented with a 5-min cartoon. The Pixar cartoon depicted a dancing lamb who became friends with a jackalope. Participants were told to pay very close attention to the video because they would be asked questions about it afterward. Immediately afterward, they were asked 18 questions about the video. Of these, 12 were answerable questions that referred to information in the video and 6 were unanswerable questions. A list of all 18 questions is provided in the Appendix. An example of an answerable question was, “How many snakes were there in the cartoon?” (there was one snake). An example of an unanswerable question was, “What was the color of the boats in the water?” (there were no boats in the water). Questions were presented to all participants in the same chronological order during both interviews with answerable and unanswerable questions intermixed. All interviews were conducted individually by the same researcher.

Participants in the voluntary confabulation condition were discouraged from speculating about answers to questions; they were encouraged to respond “don’t know” if they did not know the answer to any question. Children were told that if they thought a question was difficult and did not know how to respond, they should just respond by saying “don’t know”.

Participants in the forced confabulation condition were told to make their best speculation in answering each question even if they were unsure; they were not given a “don’t know” response option during the initial interview. They were told that they needed to respond to every question and that if a question seemed difficult to answer, they should “just guess”. When participants in the forced confabulation condition responded “don’t know”, they were twice encouraged to make their best speculation. If participants insisted that they did not know how to respond, the question was moved to the end of the interview and the experimenter later returned to the unanswered question and again encouraged the participants to respond. However, this method of forcing confabulation with reluctant participants was rarely necessary because of the ease of eliciting confabulated responses. In addition, it was predetermined that a session would be terminated if a child showed signs of distress or a lack of desire to participate. Only 1 child was terminated for this reason, and that child was excluded from the study.

Following procedures of Pezdek et al. (2007), 1 week later at Time 2 children were individually reintroduced to the same experimenter and asked the same series of questions about the target video. Participants were encouraged to think back to the cartoon they had viewed 1 week ago. At Time 2, participants in both conditions were told to say “don’t know” if they were unsure of an answer to any question (the same instructions given at Time 1 to children in the voluntary confabulation condition). It is important to note that the same interviewer was used at both Time 1 and Time 2 interviews. Although this may have increased children’s desire to appease the interviewer at Time 2, we followed this procedure to comply with the request of the school to work with only one researcher.

Data coding

All interviews were digitally recorded and later transcribed and coded for children’s responses to both answerable and unanswerable questions. Responses were coded by a second experimenter who was blind to the experimental conditions. Before data collection began, a rubric of responses was generated that described, for all questions, what responses were correct and incorrect and what responses would be considered as consistent or inconsistent from Time 1 to Time 2. The coder was

---

1 In this study, we recorded the number of times the interviewer needed to prompt children to provide a response before a confabulation occurred at Time 1 for unanswerable questions. Across all children in the forced confabulation condition who remained in the analyses, only 8 children needed to be prompted to provide a response after they initially said “don’t know” to any unanswerable question. In addition, 6 of these children needed only one additional reminder to provide a response to one unanswerable question. The remaining 2 children needed one additional prompt after the initial questioning on two questions only. The fact that it was rarely necessary to repeat or move questions to the end of the list suggests that there was a general ease of eliciting confabulated responses.
provided with this coding rubric. A lenient coding system was used such that synonyms were acceptable. For example, if the question asked, “What does the lamb like to do?” and the child initially responded during the first interview by saying “tap” but during the second interview said the lamb liked to “dance”, that was considered as consistent and coded as correct for both interviews. To ensure reliable coding, a separate research assistant coded the first 20% of the Time 1 and Time 2 interviews to verify interrater agreement. Cohen’s kappa was calculated for each question asked to determine the interrater reliability. All question responses were coded with adequate interrater reliability of at least .80.

**Results**

Responses to unanswerable and answerable questions were analyzed separately. The first analysis assessed responses to unanswerable questions at Time 1 as a function of age and condition. The second analysis assessed responses to unanswerable questions at Time 2, given that participants provided an answer to each question at Time 1, as a function of age and condition. The third analysis examined the proportion of responses at Time 1 to answerable questions in each category (correct, incorrect, “don’t know”, and correct rejections) as a function of age and condition. The fourth and final analysis examined the pattern of responses to answerable questions at Time 2, given that the correct answer to each question had been provided at Time 1, as a function of age. In each of these four analyses the pattern of responses to each response type (correct, incorrect, and “don’t know”) was assessed.

**Analyses of response rates to unanswerable questions at Time 1**

The first set of analyses assessed the frequency with which unanswerable questions were responded to by younger and older children in the forced and voluntary confabulation conditions. These data are presented in Table 1. How frequently were responses given to unanswerable questions at Time 1? To assess this question, a 2 (forced vs. voluntary confabulation condition) × 2 (6-year-olds vs. 9-year-olds) between-participants analysis of variance (ANOVA) was conducted on each of the three response types separately (i.e., provided a response, answered “don’t know”, or correct rejection of having witnessed the event in question). In none of these analyses was there a significant main effect of age or a significant age by condition interaction. As in previous research, participants in the forced confabulation condition (\(M = .94, SD = .02\)) were significantly more likely to give a response (rather than saying “don’t know”) to unanswerable questions at Time 1 than were those in the voluntary confabulation condition (\(M = .52, SD = .22\), \(F(1,91) = 128.30, p < .001, MS_e = .07, \eta^2_p = .58\). Participants in the forced confabulation condition provided responses to unanswerable questions target items nearly all of the time.

**Table 1**

Mean proportions (and standard deviations) of unanswerable questions that were responded to at Time 1 and Time 2 with a self-generated response, a “don’t know” response, or a correct rejection response (i.e., participants insisted that they did not see the proposed item) for 6-year-olds and 9-year-olds.

<table>
<thead>
<tr>
<th>Time 1</th>
<th>6-year-olds</th>
<th></th>
<th>9-year-olds</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Voluntary confabulation</td>
<td>Forcing</td>
<td>Voluntary confabulation</td>
<td>Forcing</td>
</tr>
<tr>
<td>Generated response</td>
<td>(n = 25)</td>
<td>(n = 24)</td>
<td>(n = 23)</td>
<td>(n = 23)</td>
</tr>
<tr>
<td>Don’t know</td>
<td>.54 (.25)</td>
<td>.92 (.16)</td>
<td>.51 (.19)</td>
<td>.97 (.10)</td>
</tr>
<tr>
<td>Correct rejection</td>
<td>.44 (.23)</td>
<td>.07 (.15)</td>
<td>.49 (.21)</td>
<td>.03 (.08)</td>
</tr>
<tr>
<td></td>
<td>.02 (.08)</td>
<td>.00 (.03)</td>
<td>.01 (.04)</td>
<td>.00 (.00)</td>
</tr>
<tr>
<td>Time 2</td>
<td>Generated response</td>
<td>(n = 31)</td>
<td>(n = 29)</td>
<td>(n = 24)</td>
</tr>
<tr>
<td>Don’t know</td>
<td>.58 (.31)</td>
<td>.79 (.18)</td>
<td>.61 (.24)</td>
<td>.82 (.19)</td>
</tr>
<tr>
<td>Correct rejection</td>
<td>.42 (.31)</td>
<td>.21 (.18)</td>
<td>.39 (.24)</td>
<td>.18 (.19)</td>
</tr>
<tr>
<td></td>
<td>.00 (.00)</td>
<td>.00 (.00)</td>
<td>.00 (.00)</td>
<td>.00 (.00)</td>
</tr>
</tbody>
</table>
Regarding the analysis of “don’t know” responses (see second row of Table 1), participants in the forced confabulation conditions ($M = .05, SD = .12$) were significantly less likely to say “don’t know” to unanswerable questions at Time 1 than were those in the voluntary confabulation condition ($M = .46, SD = .22$), $F(1, 91) = 134.31, p < .001, MS_e = .05, \eta^2_p = .59$. Few correct rejections (stating that an unanswerable or fabricated item/event was not present in the video) occurred in any condition (see third row of Table 1).

By examining responses to unanswerable questions at Time 1, a baseline for responding to unanswerable questions was established before examining how participants’ Time 2 responses might reflect underlying developmental differences in remembering prior confabulations.

**Analyses of responses to unanswerable questions at Time 2**

The pattern of responses to unanswerable questions at Time 2 is of primary importance in this study; it is these responses that we know were confabulated because the information relevant to answering each of these questions was not actually presented in the video. Fig. 1 depicts a graphical representation of the mean numbers of confabulated items at Time 2 by age and condition. The primary analyses address the mean proportion of unanswerable questions that were responded to at Time 2 with the same answer, a different answer, a “don’t know” response, or a correct rejection (i.e., the participant insisted at Time 2 that he or she did not remember seeing the questioned item in the video) given that the item was answered at Time 1. Thus, for these scores the number of questions included in each participant’s proportion score varied depending on whether that participant had answered the same question at Time 1. These results for unanswerable questions are presented in Table 2.

If participants generated a response to an unanswerable question at Time 1, how likely were they to recall this same answer 1 week later at Time 2, and did the likelihood of being consistent with these confabulated responses differ as a function of age and condition? To address this question, a 2 (forced vs. voluntary confabulation condition) × 2 (6-year-olds vs. 9-year-olds) between-participants ANOVA was conducted on each of the four response types indicated in Table 2. The first analysis focused on the proportion of times participants responded with the same confabulated response at Time 1 and Time 2 (see results in the first row of Table 2). Only the main effect of age was significant. The 6-year-olds ($M = .32, SD = .26$) were significantly less likely to give the same confabulated response at

![Fig. 1. Mean numbers of confabulated responses (out of six) to unanswerable questions at Time 2 as a function of condition and age. Error bars represent 95% confidence intervals.](image-url)
Time 1 and Time 2 than were the 9-year-olds (M = .65, SD = .23), F(1,91) = 27.51, p < .01, MSe = 2.67, \( \eta^2_p = .23 \).

The second ANOVA on these data assessed the proportion of Time 2 responses to unanswerable questions that were different from each participant’s Time 1 response as a function of condition and age. These data are in the second row of Table 2. The 6-year-olds (M = .47, SD = .24) were significantly more likely than the 9-year-olds (M = .24, SD = .21) to respond with a different confabulated response at Time 2 given that they initially confabulated a response at Time 1, F(1,91) = 15.40, p < .01, MSe = .98, \( \eta^2_p = .15 \). There was also a significant main effect of condition; participants in the forced confabulation condition (M = .36, SD = .24) were significantly more likely to generate a different confabulated response at Time 2 than they had initially given at Time 1 than were participants in the voluntary confabulation condition (M = .21, SD = .17), F(1,91) = 9.12, p < .01, MSe = .58, \( \eta^2_p = .09 \). The age by condition interaction was not significant.

The third ANOVA on these data assessed the proportion of responses to unanswerable questions that were responded to with a “don’t know” response at Time 2 given that a Time 1 response had been confabulated. These data are in the third row of Table 2. Only the main effect of age was significant in this analysis. The 6-year-olds (M = .30, SD = .19) were significantly more likely than the 9-year-olds (M = .16, SD = .13) to respond “don’t know” at Time 2 given that they had confabulated a Time 1 response, F(1,91) = 5.11, p < .05, MSe = .38, \( \eta^2_p = .05 \).

In the last row of Table 2 are the mean proportions of items to which participants insisted that they did not see the proposed item. These are labeled as correct rejections. Few correct rejections occurred in any condition.

### Analyses of responses to answerable questions at Time 1

The pattern of responses to answerable questions reveals response accuracy for questions regarding information that was actually presented in the video. The effect of forced confabulation on these responses is interesting because it relates to whether pressuring children to respond to questions increases their response accuracy for information that was actually presented. A 2 (forced vs. voluntary confabulation condition) \times 2 (6-year-olds vs. 9-year-olds) between-participants ANOVA was conducted on correct responses, incorrect responses, and “don’t know” responses separately. These data are presented in Table 3. The large majority of all Time 1 and Time 2 responses to answerable questions were correct, suggesting that the children in this study were attending to the video and attempting to be correct on test questions. As predicted, there was a main effect of age for correct Time 1 responses to answerable questions (see first row of Table 3); older children (M = .95, SD = .03) were significantly more likely to give a correct Time 1 response to answerable questions than were younger children (M = .89, SD = .04), F(1,91) = 11.16, p < .01, MSe = .08, \( \eta^2_p = .11 \). There was also a significant main effect of condition; children in the forced confabulation condition (M = .94, SD = .04) were significantly more likely to provide a correct response to an answerable question at Time 1 than were children in the voluntary confabulation condition (M = .88, SD = .03), F(1,91) = 12.21, p < .01, MSe = .09, \( \eta^2_p = .12 \). The age by condition interaction was not significant. There were no significant effects in
the analyses of either the incorrect responses or “don’t know” responses at Time 1; responses of these types were infrequent.

Analyses of responses to answerable questions at Time 2

The final set of analyses assessed how participants responded at Time 2 given that they answered a question correctly at Time 1. The effect of forced confabulation on these responses is interesting because it relates to whether pressuring children to respond to questions increases their response accuracy for information that was actually presented after a time delay in a follow-up interview. Three 2 (forced vs. voluntary confabulation condition) × 2 (6-year-olds vs. 9-year-olds) between-participants ANOVAs were conducted to assess whether participants were equally likely to give a correct response, an incorrect response, or a “don’t know” response 1 week later at Time 2. These data are presented in Table 4. The first set of analyses focused on correct responses only (see the first row of Table 4). Older children ( \( M = .94, SD = .08 \)) were significantly more likely to provide a correct response at Time 2 than were younger children ( \( M = .88, SD = .17 \)), \( F(1,91) = 5.90, p < .05, MS_e = .11, \eta^2_p = .06 \). There was a main effect of condition; children in the forced confabulation condition ( \( M = .95, SD = .06 \)) were significantly more likely to give correct responses to answerable questions at Time 2 than were those in the voluntary confabulation condition ( \( M = .87, SD = .18 \)), \( F(1,91) = 7.43, p < .01, MS_e = .13, \eta^2_p = .07 \). The age by condition interaction was not significant.

The second ANOVA on these data assessed the proportion of Time 2 responses to answerable questions that were incorrect given that a participant responded correctly at Time 1. These data are in the second row of Table 4. Although more correct responses were generated at Time 2 for participants in the forced confabulation condition than for those in the voluntary confabulation condition, participants in the forced confabulation condition ( \( M = .04, SD = .05 \)) were not significantly more likely to report incorrect responses to answerable questions at Time 2 than participants in the voluntary confabulation condition ( \( M = .05, SD = .07 \)), \( F(1,91) = 2.30, p > .05, MS_e = .01, \eta^2_p = .02 \). Inconsistent with previous findings with adults (Pezdek et al., 2007), forcing children to speculate improved their correct responses at Time 1 without significantly raising their likelihood of giving an incorrect response at

Table 3
Mean proportions (and standard deviations) of answerable questions that were responded to at Time 1 with a correct answer, an incorrect answer, or a “don’t know” response.

<table>
<thead>
<tr>
<th>Time 1 response</th>
<th>6-year-olds</th>
<th>9-year-olds</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Voluntary confabulation (n = 25)</td>
<td>Forced confabulation (n = 24)</td>
</tr>
<tr>
<td>Correct answer</td>
<td>.86 (.10)</td>
<td>.92 (.08)</td>
</tr>
<tr>
<td>Incorrect answer</td>
<td>.07 (.07)</td>
<td>.08 (.08)</td>
</tr>
<tr>
<td>Don’t know</td>
<td>.07 (.07)</td>
<td>.00 (.01)</td>
</tr>
</tbody>
</table>

Table 4
Mean proportions (and standard deviations) of answerable questions that were responded to at Time 2 with a correct answer, an incorrect answer, or a “don’t know” response given that the question was answered correctly at Time 1 for 6-year-olds and 9-year-olds.

<table>
<thead>
<tr>
<th>Time 2 response</th>
<th>6-year-olds</th>
<th>9-year-olds</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Voluntary confabulation (n = 25)</td>
<td>Forced confabulation (n = 24)</td>
</tr>
<tr>
<td>Correct answer</td>
<td>.83 (.23)</td>
<td>.93 (.07)</td>
</tr>
<tr>
<td>Incorrect answer</td>
<td>.05 (.07)</td>
<td>.04 (.06)</td>
</tr>
<tr>
<td>Don’t know</td>
<td>.12 (.23)</td>
<td>.03 (.04)</td>
</tr>
</tbody>
</table>
Time 2. The main effects of age and the age by condition interaction were not significant in this analysis.

The third ANOVA on these data assessed the proportion of responses to answerable questions that were responded to with a “don’t know” response at Time 2 given that a correct response was given at Time 1. These data are presented in the third row of Table 4. There was a main effect of age: the 6-year-olds ($M = .02$) were significantly less likely to respond “don’t know” at Time 2 after giving a correct response at Time 1 than were the 9-year-olds ($M = .07$), $F(1,91) = 4.93$, $p < .05$, $MSE = .08$, $\eta^2_p = .05$. There was also a main effect of condition; given that a correct response was initially provided at Time 1, children in the forced confabulation condition ($M = .02$, $SD = .04$) were significantly less likely to respond “don’t know” at Time 2 than were those in the voluntary confabulation condition ($M = .07$, $SD = .18$), $F(1,91) = 4.57$, $p < .05$, $MSE = .07$, $\eta^2_p = .05$. There was no significant interaction in this analysis of “don’t know” responses.

**Discussion**

The major finding in this study was that although 6-year-olds and 9-year-olds were equally likely to provide a response to an unanswerable question at Time 1, at Time 2 the 9-year-olds were significantly more likely than the 6-year-olds to unwittingly repeat their Time 1 confabulated responses. This age effect was consistent in both the voluntary and forced confabulation conditions; there was no significant age by condition interaction in this analysis.

The findings suggest that forced confabulation is likely to impair the eyewitness memory of both 9-year-olds and 6-year-olds but in different ways. Two sets of findings support this conclusion. First, as shown in Table 1, the data for children in the forced confabulation condition show that 79% of 6-year-olds and 82% of 9-year-olds gave an incorrect (“generated”) response to the unanswerable questions at Time 2. In other words, children in both age groups showed equally “impaired eyewitness memory” at Time 2. The evidence does not indicate that in general memory is significantly more impaired for 9-year-olds than for 6-year-olds. However, as shown in Table 2, there is a difference between the age groups concerning the type of impairment. Among 9-year-olds, memory impairment was relatively more likely to take the form of repeating the same (erroneous) answer at Time 2 as at Time 1. Among 6-year-olds, however, there was a greater tendency to generate a “new” error at Time 2 that was different from the error at Time 1. Whereas 6-year-olds in the forced confabulation condition gave consistent Time 1 to Time 2 responses to unanswerable questions only 32% of the time, 9-year-olds did so 58% of the time. These findings suggest age differences in memory for the information that had been initially confabulated.

The results of this study are consistent with the two-stage model of forced confabulation proposed by Pezdek et al. (2007) and extend this model to account for developmental differences in the underlying cognitive processes. That is, forced confabulation requires that, first, an individual must recall at Time 2 his or her self-generated answer to the unanswerable question from Time 1 and, second, the individual must make a source monitoring error at Time 2 and confuse the self-generated answer with information actually observed. Results of previous studies have demonstrated age differences in the second stage of this model, susceptibility to source monitoring errors (Foley & Johnson, 1985; Foley et al., 1983; Parker, 1995), in the direction that younger children are more susceptible to these errors than are older children. The results of this study suggest that age differences during the first stage of this process—the ability to recall at Time 2 one’s self-generated answer from Time 1—appear to account for developmental differences in the persistence of confabulated information in memory as well. In addition, because in this study 9-year-olds were more likely than 6-year-olds to remember and repeat their Time 1 response at Time 2, this suggests that this first process, the memory process, has a stronger role in children’s forced confabulations than does the second process, source monitoring. This is an important contribution of this study to the research on developmental differences in forced confabulation.

There are significant developmental differences in forced confabulation reported between previous studies with adults and the current study with children. Pezdek et al. (2007) reported that information generated from forced confabulation was less likely to be remembered than information voluntarily
self-generated. Furthermore, Pezdek, Lam, and Sperry (2009) reported that if an answer to an unanswerable question was forced confabulated at Time 1, that answer was more likely to be repeated at Time 2 if it had been other-generated (suggested in the Time 1 question) rather than self-generated (fabricated by the participant). From the framework of signal detection theory, we know that for adults forced confabulation serves to produce less cautious responses. That is, forced confabulation lowers one's response criterion (i.e., reduces beta), such that on average self-generated items are considered to be of greater strength in the voluntary confabulation condition than in the forced confabulation condition (Gombos et al., 2012). It is interesting to note that in the current study there were no interactions of age by confabulation condition. The three nonsignificant effects of condition (i.e., on frequency of same response, “don’t know” response, and correct rejection) in the current study suggest that there is less likely to be a difference in memory strength between items confabulated in the voluntary confabulation condition and those confabulated in the forced confabulation condition for children. In this study, the memory strength of self-generated confabulated items to both forced and voluntary self-generated responses was high.

It is also of interest that children in the forced confabulation condition provided more correct responses to answerable questions than did children in the voluntary confabulation condition. This finding is inconsistent with results with adults reported by Pezdek et al. (2007). In that study, encouraging participants to speculate significantly increased their tendency to give both correct and incorrect responses to answerable questions. The current findings suggest that although forced confabulation may be detrimental to children's memory for unanswerable questions, it may improve their correct responses to answerable questions, making them no more likely to report incorrect responses for answerable questions. Further research is needed to assess why forced confabulation might positively influence children's accuracy in answering answerable questions without also increasing their rate of incorrect responding. However, it is important to note that in nonexperimental situations, it is rarely clear which questions are answerable and which are unanswerable for an individual, and being forced to guess may increase the likelihood of confusing self-generated information that did not occur with information that was actually witnessed.

There are several caveats to consider. First, in forensic interviews, it is probably more typical that negative events, such as abuse, are discussed. In the current study, the target event was not a negative event but rather a cartoon video. Future research might investigate whether differences arise in children's tendencies to confabulate and recall confabulations with positive versus negative events. Second, different effects might be expected if a different ratio of unanswerable to answerable questions was used. For example, if there was only one unanswerable question, children might feel more confident in responding “not present” (in the video), having already demonstrated their knowledge of the event by responding to mostly answerable questions. On the other hand, with more unanswerable questions, children might feel less confident in their memory and, thus, more obliged to confabulate. Third, there was a near floor correct rejection rate. However, this floor effect does not detract from our main findings regarding rates of children's confabulated responses over time. Finally, an alternative explanation of children's confabulations is that participants were able to remember that they initially confabulated their response, and in attempts to appease the interviewer and appear consistent they simply repeated this response at Time 2 even though they were aware of the fabrication. However, this explanation seems unlikely, especially for younger participants. The Time 2 confabulated responses reported by 6-year-olds were often inconsistent with their Time 1 confabulations, suggesting that the children did not merely repeat their original response to appease the interviewer.

This research is forensically relevant because it is common for children to be repeatedly interviewed, sometimes over long periods of time (Goodman, Taub, Jones, & England, 1992; Malloy, Lyon, & Quas, 2007; Poole & White, 1993). In these repeated interviews, an investigating officer may press a child witness to answer questions because the officer thinks the child may know the answers but simply is reluctant to respond. Although the findings of this study indicate that forcing children to speculate might not increase their incorrect responses to answerable questions, forcing speculation may result in an increase in confabulated responses that are later reported to have been part of the target event viewed. In the real world, the investigating officer does not know for a particular witness which questions are answerable (i.e., the witness did observe information relevant to answering the question) and which are unanswerable (i.e., the witness did not observe information relevant to answering
the question). The results of this study suggest that pressing child witnesses to answer questions about which they are reluctant may be an ineffective practice; doing so increases the rate of confabulating answers and increases the probability that such confabulated answers will be repeated over time, especially for older children. This study provides a situation in which the consistency of children’s responses over time is not an indication of the accuracy of their eyewitness memory. Additional research on developmental differences in forced confabulation will likely lead researchers and interviewers to understand the conditions that will enhance the accuracy of children’s eyewitness memory and reduce the deleterious effects of suggestive influences such as forced confabulation.

Acknowledgments

This study was conducted as the M.A. thesis of Stacia Stolzenberg under the supervision of Kathy Pezdek. We thank Bernadette Boyle (principal), and the teachers and students at Our Lady of the Assumption School in Claremont, California, for their participation in this research. We are also grateful to Tiffany Berry for her comments on an earlier draft of the article and to Jaya Roy for her assistance with data coding. This research was supported by a grant from the Fletcher Jones Foundation to Kathy Pezdek.

Appendix A

List of 18 questions asked of participants. Questions asked of participants with confabulated items are presented in bold:

- How many fish were swimming at the beginning of the video?
- Where do the animals live?
- **What color was the boat in the river?**
- How many teeth does the lamb have sticking out of his mouth?
- How many snakes were there in the video?
- What does the lamb like to do?
- **What clothes does the frog have on?**
- What happened to the lamb that made him sad?
- **What color was the shirt worn by the man in the wagon?**
- What did the animals do to the sheep that made him feel bad?
- What color was the lamb’s skin?
- What color were the lamb’s eyes?
- What does the jackelope tell the lamb to make him happy?
- **What does the jackelope say to the gophers or brown animals in the holes?**
- How many owls were in the video?
- **What color were the flowers on the plants behind the lamb?**
- How often does the lamb get his hair cut?
- **Describe the hat that the jackelope was wearing.**

References


