APPLIED COGNITIVE PSYCHOLOGY
Appl. Cognit. Psychol. **15**: 353–371 (2001)
DOI: 10.1002/acp.708

Recounting the Same Events Again and Again: Children's Consistency Across Multiple Interviews

CAROLE PETERSON,* LISA MOORES and GINA WHITE

Memorial University of Newfoundland, Canada

SUMMARY

Children (2–13 years at time of injury) were interviewed four times about an injury that required hospital Emergency Room treatment, namely at 1 week, 6 months, 1 year, and 2 years. The consistency of children's reports was assessed and all children gave mostly the same information at each interview, although consistency was higher for older children and for injury rather than hospital details. Furthermore, details recalled at every interview were virtually always accurate while details that were sometimes omitted were a little less likely to be accurate. New information that was introduced after 6 months was more likely to be accurate than inaccurate but new information introduced at 1 or 2 years post-injury was just as likely to be wrong as right (except for 12–13-year-olds). Implications for forensic situations are discussed. Copyright © 2001 John Wiley & Sons, Ltd.

Children who are witnesses or victims in forensic situations are subsequently interviewed multiple times by multiple people. The consistency of their accounts across different interviews can be an important issue in whether they are believed by police, judges and jurors. On the one hand, inconsistency can seriously undermine the credibility of children's testimony, leading to discrediting of their accounts, and it has been found to have a substantial negative impact on jurors (Leippe *et al.*, 1991; Ross *et al.*, 1987). According to Steward *et al.* (1996, pp. 6–7), 'the consistency of a child's report was rated in a recent national survey as one of the top three criteria that professionals use to assess the accuracy of allegations of child abuse. On the other hand, some courts have ignored such inconsistency because of the child's young age. For example, in a murder case described by Poole and White (1995), the 5-year-old witness identified more than a dozen perpetrators and four different murder weapons across multiple interviews; nevertheless, her testimony led to convictions.

Likewise, the same sort of ambiguity and disagreement exists when the courts are confronted with new information which is added in subsequent interviews that was not present in earlier interviews, a phenomenon often termed reminiscence. To quote Poole

*Correspondence to: Carole Peterson, Psychology Department, Memorial University of Newfoundland, St Johns, Newfoundland, Canada AIB 3X9. E-mail: carole@play.psych.mun.ca

Contract grant sponsor: Natural Sciences and Engineering Research Council of Canada.
Contract grant number: OGP0000513.
Contract grant sponsor: Memorial University Undergraduate Career Experience Program.

and White, (1995, p. 33), 'According to Cassell, an experienced attorney, reminiscence can be used to discredit a child on the basis that the new facts must have been "planted" by an intervening event'. In contrast, theoretical accounts suggest that repeated interviews can aid memory and that such interview repetition may be an effective strategy for eliciting additional accurate information (Fivush *et al.*, in press). In fact, several studies have suggested that repeated interviewing can increase the amount of information reported (Brainerd and Ornstein, 1991; Dent, 1991; Goodman and Clarke-Stewart, 1991; Howe *et al.*, 1992), and some of the additional interviews given to children who are involved in forensic situations are carried out specifically in attempts to elicit additional information (Poole and White, 1995).

The research reported here addresses both of the issues raised above. It explores the consistency of information provided by children across multiple interviews as well as the accuracy of consistently recalled details. Each interview is also compared with prior interviews to identify information that is new to a later interview. The accuracy of that new information is then compared with the accuracy of previously recalled, 'old' information.

Several reports have focused on either children's consistency across multiple interviews or the accuracy of information that is new to later interviews (or both). Turning first to investigations of consistency, some recent reviews have appeared (Fivush *et al.*, in press; Fivush and Schwarzmueller, 1995; Poole and White, 1995). The type of interview seems to have important effects. When Fivush, Hudson, and their colleagues (Fivush and Hamond, 1990; Fivush and Shukat, 1995; Hudson, 1990; Hudson and Fivush, 1991) asked preschool-aged children to tell them about various events such as visits to museums, they found that up to 80% of the information provided during later interviews about the same events was different, although mostly accurate.

Importantly, the interviews themselves differed; there was no core of similar or identical questions that were asked on multiple occasions. Instead, questions differed considerably across interviews. As well, children are sensitive to the listener's state of knowledge and often do not repeat information the listener already knows (Fivush and Schwarzmueller, 1995; Menig-Peterson, 1975). Since many of the repeated interviews in the above studies were conducted by interviewers familiar with the events being recounted, this may have negatively affected consistency. The age of the children may also play a role since school-aged children showed considerably more consistency than did preschoolers (Hudson and Fivush, 1991).

What happens when the same questions are asked repeatedly across interviews? We will not discuss repetition of the same questions within the same interview since these data have recently been reviewed (Fivush *et al.*, in press; Fivush and Schwarzmueller, 1995; Poole and White, 1995) and all authors concur that this intra-interview repetition has a serious negative impact on the consistency of children's reports. There have been mixed reports of children's consistency across interviews that ask the same questions. In studies where delays across interviews are short (e.g. a week), consistency can be high (Poole, 1995; Schwarzmueller, 1997, unpublished manuscript), although Poole and White (1991) found that 4-year-olds inconsistently responded to yes/no questions. However, preschoolers' responses to yes/no questions seem particularly likely to be inaccurate (Peterson and Biggs, 1997; Peterson *et al.*, 1999) and so question format may have affected Poole and White's results. As well, the same questions were repeated within as well as across interviews, a practice that jeopardizes children's response accuracy (Fivush *et al.*, in press; Fivush and Schwarzmueller, 1995; Poole and White, 1995) and may have undermined cross-session consistency.

Few studies have addressed consistency across interviews that are widely spaced in time. Poole and White (1993) found quite low consistency between interviews on questions answered initially and 2 years later by children who were 4 years old initially. However, they exclusively analysed yes/no responses—a question format that is problematic. In two other studies, investigators looked at the consistency of children's recall of medical or quasi-medical events. In a study of preschoolers' recall of pediatric exams, Steward *et al.* (1996) interviewed children shortly afterwards as well as 1 and 6 months later. At 1 month, only 51% of information was the same as in the first interview, and at 6 months, only 52% was present in at least one of the previous interviews (but only 25% was present in both). Information that was consistently present across interviews was highly accurate, however. In another study, Salmon and Pipe (1997) had 3- and 5-year-old children take part in a quasi-medical event in which a 'sick' teddy bear was examined. When children were verbally interviewed 3 days and 1 year later, only 15% and 25% of the items mentioned initially by the younger and older children respectively were repeated in the 1-year interview. These percentages were higher when props were also used—but still only 45% and 53% for younger and older children, respectively. Parallel to Steward *et al.* (1996), information that was consistently mentioned by the children across interviews was highly accurate. Overall, even though both Salmon and Pipe (1997) and Steward *et al.* (1996) found that repeated information was almost always accurate, they nevertheless reported that children provided a considerable amount of different information in subsequent interviews.

In summary, reports of children across interviews were inconsistent. The degree of consistency varied considerably depending upon interview questions; if questions differed across interviews, children reported quite different information but if questions were the same, they tended to be less inconsistent. As well, yes/no questions were especially inconsistently responded to by preschoolers. The only studies showing high consistency had a very short delay between interviews, with longer delays associated with greater inconsistency. However, the two studies with long delays (and who did not use only yes/no questions) were both medical check-ups (of self or a teddy bear)—events that are seldom unique in a child's experience.

It would be instructive to investigate children's consistency when recalling other sorts of events, and specifically ones that are more unique since uniqueness seems to be an important property that fosters excellent long-term recall (see review in Howe, 1997). In the present study, children's injuries that occurred as a result of accidents are investigated. These injuries were serious enough that the children had to be taken to a hospital Emergency Room for medical treatment, thus, the events were not only unique, they were also personally relevant and highly salient—properties that seem to foster better long-term recall (Howe, 1997).

The accuracy of new information that is added in subsequent interviews is also an important forensic issue. Laboratory studies find that new information is likely to be accurate (e.g. Howe *et al.*, 1992). Likewise, Fivush, Hudson, and their colleagues found that in the free-ranging interviews they studied, new information was mostly accurate (Fivush and Hamond, 1990; Fivush and Shukat, 1995; Hudson, 1990; Hudson and Fivush, 1991). However, very different conclusions were reached by Steward *et al.* (1996) and Salmon and Pipe (1997). Steward and her colleagues found that only about 60% of new information reported in either the 1- to 6-month repeated interviews was accurate, while Salmon and Pipe found that only about half of new information was accurate, with younger children having poorer accuracy rates than older children. Thus, in contrast to

Fivush and Hudson, both of these research teams concluded that information that is repeated is very likely to be accurate whereas information that is included for the first time in later interviews is highly suspect because it is so frequently wrong.

It is difficult to reconcile the divergent findings reported above. Some investigators suggest that laboratory studies of stimuli such as word lists are not necessarily comparable to reports of real-world events (Fivush *et al.*, in press; Fivush and Schwarzmueller, 1995; Poole and White, 1995). It is also relatively easy to account for the high accuracy rates of the new information provided by the children in Fivush and Hudson's work since the interviews often asked quite different questions. Therefore it is unclear how much new information might have been provided had the same questions been asked. There are also potentially ambiguating factors in both Steward *et al.*'s (1996) and Salmon and Pipe's (1997) studies: both studies were investigations of children's memory for medical exams. Since children typically have considerable experience with doctor check-ups (both real and play), they may well develop scripts for typical doctor actions or misremember which components they had experienced during medical check-ups were actually part of any specific medical exam. Confusions between similar-but-different events have frequently been found (e.g. Hudson and Nelson, 1986), and Salmon and Pipe (1997) suggest that when children introduce new infromation after a delay, it may relate to different-but-similar events rather than the target event. Thus, it is important to extend research to other events besides check-ups.

In the current study, children were interviewed about the same events (in which they suffered personal injury) on four different occasions over a 2-year period. The same interview was used each time, although the interview was conducted by different people. Consequently, this study can address issues of consistency over time as well as the accuracy of information that gets added in later interviews in comparison with information that is repeated from prior interviews. The ages of the children spanned the years from 2 to 13 when the target events occurred, and thus age comparisons can be made across a wide age range.

Other data on the children in the current study have been reported elsewhere (Peterson, 1996, 1999; Peterson and Bell, 1996). However, the data are scored differently here in order to answer different questions. In previous research, children's transcripts were scored in terms of whether previously selected prototype elements were present or absent, i.e. there was an all-or-nothing score for each element. Such a scoring system allows one to assess which elements of a prototypical experience children recall and which they forget over time. However, in the present study the children's responses are analysed in more detail, and with a focus on consistency.

It is hypothesized that consistency will increase with age, with 2-year-olds being particularly inconsistent, in keeping with previous reports (Peterson, 1999; Peterson and Bell, 1996). It is also expected that information that is repeated across interviews will be highly accurate, whereas new information that is introduced for the first time in later interviews will be less so. In addition, the details surrounding the children's injury event are more unique than many of the details surrounding their hospital treatment, since all children in the community in which the study was conducted are taken to the same hospital Emergency Room (ER) for every major illness as well as for most injuries; it is also the only place children can be taken for treatment when doctors' offices are closed (i.e. evenings, weekends, and holidays). Thus, a number of things that happened in the ER at the time of the target event had been experienced by the children on multiple other occasions. (See Peterson and Bell, 1996, for data on children's prior experiences with this ER.) Therefore, it is hypothesized that children will be more consistent as well as more accurate in their recall of injury details than of hospital details.

METHOD

Participants

The 96 children (45 girls and 51 boys) in this study were recruited from the ER of the only children's hospital in Newfoundland, Canada. They were mostly White from mixed SES backgrounds. They had experienced a trauma injury that was treated in an out-patient manner, including lacerations requiring suturing (52 subjects), bone fractures (31 subjects), burns (1 subject), dog bites (3 subjects), crushed fingers requiring drainage (2 subjects), and bandages (7 subjects). In other reports, 2-year-olds were substantially different from 3-year-olds (Peterson, 1996; Peterson and Bell, 1996); they recalled much less and made considerably more errors than did children just a year older. Because of this, the 2-year-olds are in a separate age group in all analyses, whereas combining of ages occurs among the older children. The five age groups are as follows: (1) eleven 2-year-olds (mean age at injury 2–5, range 2–2 to 2–11), (2) eighteen 3–4-year-olds (mean age at injury 3–9, range 3–2 to 4–9), (3) twenty-six 5–6-year-olds (mean age at injury 5–9, range 5–1 to 6–11), (4) twenty-two 8–9-year-olds, (mean age at injury 8–9, range 8–0 to 9–11), and (5) nineteen 12–13-year-olds, (mean age at injury 12–7, range 12–0 to 13–6).

Procedure

As described in detail in Peterson and Bell (1996), parents and children were recruited in the hospital ER and initial visits made to their homes. At this time the children were interviewed about what they recalled about their injury and subsequent treatment; as well, the parents or, if necessary, other adult witnesses were interviewed in order to provide a standard that could be used against which we could evaluate the children's information as accurate or not.

The first interview took place within a few days of the injury (mean delay = 7.3 days, range 1–20 days). Subsequent interviews took place at 6 months (mean delay 6 months 3 days, range 5–8 months), 1 year (mean delay 12 months 2 days, range 10 to 14 months), and 2 years (mean delay 24 months 6 days, range 20 to 28 months). The 2-year follow-up was unexpected. When telephone contact was made for each visit after the initial one, the interviewer asked that parents not rehearse the events with children prior to her visit.

The format of each interview was the same: free recall ('Tell me about when you hurt yourself. What happened?' 'Tell me about when you went to the hospital. What happened?') followed by probed recall using wh- questions ('Where were you when it happened? Who was with you? What did you do when you first got hurt?'). For return visits months later, free recall probes reminded the children of the target injury ('Remember that time when you broke your arm? Tell me about it. What happened?'). Yes/no questions were avoided as much as possible. The questionnaire was the same for each interview, regardless of delay. The interviews were audio-recorded and transcribed verbatim, and all scoring was done from these transcripts.

Coding

Several measures of consistency and accuracy were obtained. First, the details provided by the children in all interviews were listed on scoresheets. Then the transcript of each subsequent interview was compared with each child's previous interview(s), to determine whether each item of information was 'old', i.e. previously stated by the child, or 'new', i.e. mentioned for the first time. As well, each item was compared with the information provided by adult witnesses to determine accuracy. Consequently, the mean number and percentages of both 'old' and 'new' information in each interview that was accurate was derived.

The consistency of the children's recollection of each detail was also classified. This was done in two ways to reflect different aspects of the data, namely, both a mean *consistency* score and a mean *inconsistency* score were derived. Both scores were assigned only for details that were recalled at least twice, since no score could be assigned if the child recalled a detail only once. For the *consistency* scores, a score of 6 was assigned if a child recalled the same detail the same way in all four interviews, a score of 5 if the detail was recalled the same way in three interviews and omitted in the remaining one, a score of 4 if the detail was recalled the same way in two interviews and omitted in the other two, a score of 3 if the child was consistent in three of the four interviews in which the detail was recalled (with the detail recalled differently in one interview), a score of 2 if the child was consistent in two of the three interviews in which it was recalled (and recalled it differently in one), and a score of 1 indicated greatest inconsistency. The consistency scores for all details were then averaged to obtain a mean consistency score for each child. Note, however, that children could have consistently recalled any detail correctly or incorrectly—this scoring only looked at consistency of recall, not correctness of recall.

The *inconsistency* scores were assigned as follows: a score of 3 was assigned if the child was never inconsistent, a score of 2 was assigned if a child had only one inconsistency in the recall of a specific detail, i.e. recalled it two or three times the same way and recalled the detail differently in one interview, and a score of 1 was assigned if the child had multiple inconsistencies in recall, i.e. recalled it differently each time it was recalled (whether in two, three, or all four interviews), or recalled it twice one way and twice another way. To obtain inter-scorer reliability, a second research assistant scored 20% of the transcripts; the reliability for scoring the presence and accuracy of each element averaged 98% (scored as the number of agreements divided by the number of agreements plus disagreements, range 83% to 100%).

RESULTS

There were two main foci for this study: consistency of children's responses across four interviews, and the accuracy of information that was recalled for the first time in later interviews. To assess consistency of recall, three measures were analysed: the percentage of details in each interview that was 'old' information was assessed, a mean consistency score was derived for each child, and a mean inconsistency score was derived for each child. In addition, to see whether consistent recall was highly associated with accurate recall, the accuracy of details that were consistently recalled was tabulated. The second focus of this study was to assess the recall accuracy of items that were recalled for the first time in later interviews (i.e. 'new' information). To do this, the accuracy of 'new' items were calculated and compared to the accuracy of 'old' items. The significance level for all analyses is $p < 0.05$.

Consistency of recall

Three interviews took place after the initial one. Table 1 presents the percentage of information in each of later interviews that is repeated from a previous interview, i.e. 'old' information. For comparative purposes, the total number of details recalled by the children is also presented in Table 1 although not analysed, since it is not the focus of this study (see Peterson, 1999). To analyse the percentage of 'old' details recalled in each interview, an ANOVA was calculated with Age (5 levels) and Gender (2 levels) the between-subjects factors and both Time (3 levels: 6 month, 1 year, and 2 year) and Event (2 levels: injury versus hospital) the within-subjects factors. Children were recalling a larger proportion of 'old' information in subsequent interviews as they got older, $F(4,86) = 16.71$, $MS_e = 357.66$. The means of the five age groups are as follows for the 2-year-olds through 12–13-year-olds, respectively: 72.5%, 81.1%, 84.1%, 91.8% and 94.6%. Paired comparisons showed that the 2-year-olds included significantly less 'old' information in their subsequent interviews than did all older age groups. The 3–4- and 5–6-year-olds did not differ from each other, nor the 8–9- and 12–13-year-olds, but these two younger groups differed from the two oldest groups. The time of the interview was also significant, $F(2,172) = 26.03$, $MS_e = 174.16$. The means for the three interview times are 80.3%, 89.4%, and 88.5% for the 6-month, 1-year, and 2-year interviews, respectively. Paired comparisons showed that the 6-month interview had less 'old' information than did either of the subsequent interviews, which did not differ from each other. Children also included more 'old' information in their recollections of their injury ($M = 88.8\%$) than of their hospital treatment ($M = 83.3\%$), $F(1,86) = 16.78$, $MS_e = 194.44$. Gender was not a significant factor.

The only significant interaction was Time × Event, $F(2,172) = 4.41$, $MS_e = 121.06$. This interaction is shown in Figure 1. Analyses of simple effects showed that at each interview time, more 'old' information was recalled about the injury event than about the hospital event. Furthermore, looking at each event separately, both showed a time effect. Overall, children presented considerably less 'old' information about their hospital treatment than their injury in the 6-month interview, but differences between events were not as great in later interviews.

The above analysis compared all four interviews. That is, an item was defined as 'old' if it was repeated from any of the prior interviews. To see how similar successive interviews were to each other, a further analysis was done on pairs of adjacent interviews. That is, the 6-month, 1-year, and 2-year interviews were separately compared with each's immediately preceding interview. An item was defined as 'old' only if it had been mentioned in the directly preceding interview. The percentage of 'old' information was 80.6%, 72.9%, and 70.6% for the 6-month, 1-year, and 2-year interviews, respectively. Comparisons of adjacent interviews were done by means of ANOVAs on the percentage of information in the later interview that was repeated from the earlier interview, with Age and Gender the between-subjects factors and Event the within-subject factor. In all three analyses the factors of Age and Event were significant, parallel to the analysis that combined all interviews. Additional significant relationships were found for the 1-year and 2-year analyses, however. For the 1-year analysis, there were significant Age × Gender and

Table 1. The mean number of details and percentage of details in each interview that are repeated (i.e. 'old' information) for both the injury and hospital event for each age group

	Time of interview											
	6-month interview				1-year interview				2-year interview			
	Injury		Hospital		Injury		Hospital		Injury		Hospital	
Age	*M*	% old	*M*	% old	*M*	% old	*M*	% old	*M*	% old	*M*	% old
2 years	7.2	69.7	3.7	59.1	10.3	83.0	4.7	72.8	10.6	78.2	5.1	74.5
(SD)	*(3.5)*	*(25.3)*	*(2.4)*	*(33.3)*	*(2.7)*	*(14.7)*	*(3.4)*	*(21.9)*	*(4.4)*	*(12.5)*	*(2.5)*	*(27.2)*
3–4 years	13.6	78.3	6.8	74.3	15.3	88.6	8.0	83.0	16.5	86.8	8.9	82.1
(SD)	*(4.1)*	*(15.8)*	*(2.8)*	*(23.9)*	*(3.8)*	*(6.0)*	*(3.1)*	*(16.5)*	*(3.5)*	*(11.5)*	*(2.9)*	*(11.9)*
5–6 years	15.1	83.3	7.5	72.4	15.9	88.5	9.6	88.5	17.9	89.1	10.3	86.1
(SD)	*(3.3)*	*(12.3)*	*(2.8)*	*(16.7)*	*(2.9)*	*(8.8)*	*(2.6)*	*(13.0)*	*(3.5)*	*(8.6)*	*(2.8)*	*(12.1)*
8–9 years	16.2	94.3	7.6	77.3	16.4	95.3	9.0	95.1	17.2	94.1	8.7	93.5
(SD)	*(2.8)*	*(8.0)*	*(2.8)*	*(18.0)*	*(3.3)*	*(8.4)*	*(2.8)*	*(7.0)*	*(3.0)*	*(6.2)*	*(2.5)*	*(8.9)*
12–13 years	13.9	94.4	9.3	88.7	14.7	98.8	10.2	90.9	13.9	99.4	9.8	95.2
(SD)	*(3.2)*	*(9.5)*	*(3.1)*	*(15.9)*	*(2.8)*	*(2.3)*	*(2.2)*	*(7.7)*	*(2.3)*	*(2.3)*	*(2.0)*	*(9.0)*
All ages	13.9	85.6	7.3	75.6	15.0	91.5	8.7	87.7	15.9	90.7	9.0	87.6
(SD)	*(4.2)*	*(16.1)*	*(3.2)*	*(22.0)*	*(3.6)*	*(9.7)*	*(3.2)*	*(14.5)*	*(4.0)*	*(10.4)*	*(3.0)*	*(14.9)*

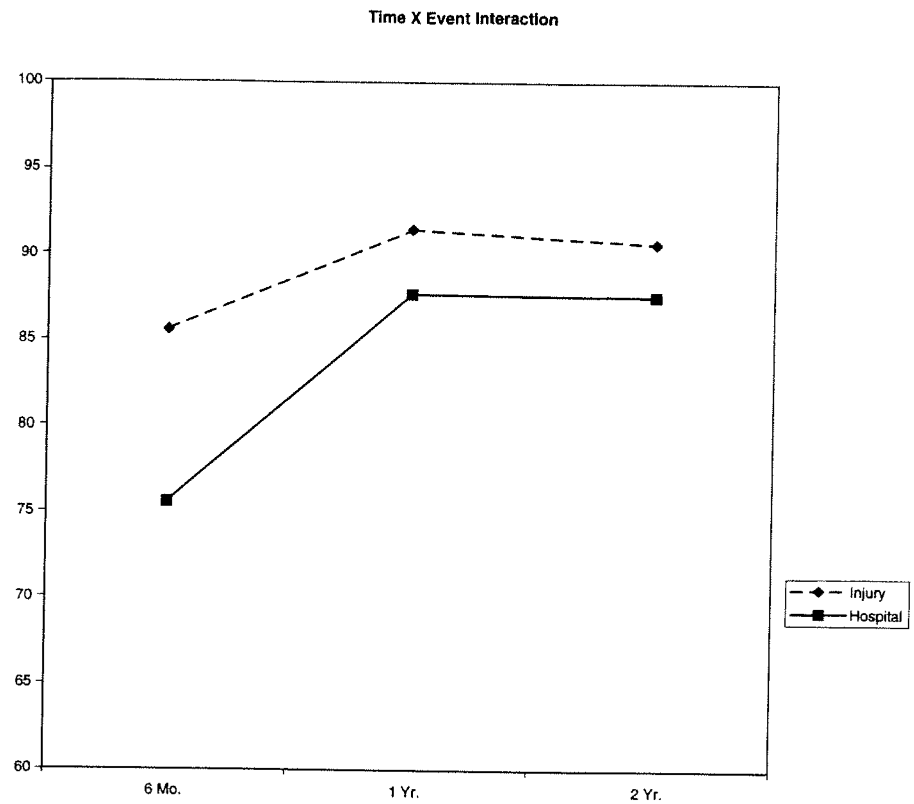

Figure 1. The percentage of 'old', previously recalled information about their injury and hospital treatment by children at interview delays of 6 months, 1 year, and 2 years

Age × Gender × Event interactions, $F(4,86) = 3.25$, $MS_e = 311.7$ and $F(4,86) = 2.73$, $MS_e = 177.7$. These interactions were due to 3–4-year-old boys repeating less information from their 6-month interview than did girls in their 1-year interview, especially about the injury event, as well as 12–13-year-old girls repeating less information about the injury event than did boys. For the 2-year analysis, there was an additional main effect of Gender, $F(1,86) = 8.09$, $MS_e = 275.8$, as well as interactions of Gender × Age, Gender × Event, and Age × Event, $F(4,86) = 3.20$, $MS_e = 275.8$, $F(1,86) = 8.23$, $MS_e = 163.4$, and $F(4,86) = 3.67$, $MS_e = 163.2$. The significant main effect for gender as well as all of the interactions were attributable to the fact that in the 2-year interview, the 2-year-old boys repeated only 7% of the information that they had provided in their 1-year interview about the hospital event—a number that was enormously different from how much of the injury event they repeated (53%), how much 2-year-old girls repeated of the hospital event (45%), and how much all older children recalled about both events.

Although the children primarily recalled 'old' information, explicit assessment of how consistently they provided the same details over time is important. Analyses of both the mean *consistency* scores and mean *inconsistency* scores were done. Recall that the values of the *consistency* score ranged from 1 (not at all consistent) to 6 (recollected consistently

Table 2. Mean consistency scores and inconsistency scores for children's recall of both injury and hospital details, for each age group and gender[a]

	Type of score							
	Consistency score				Inconsistency score			
	Injury		Hospital		Injury		Hospital	
Group	*M*	*(SD)*	*M*	*(SD)*	*M*	*(SD)*	*M*	*(SD)*
2-year girls	3.8	*(0.9)*	3.2	*(0.8)*	2.4	*(0.3)*	2.3	*(0.3)*
2-year boys	3.4	*(0.7)*	2.6	*(0.6)*	2.2	*(0.2)*	2.0	*(0.4)*
3–4-year girls	4.7	*(0.2)*	4.8	*(0.5)*	2.6	*(0.1)*	2.7	*(0.2)*
3–4-year boys	4.5	*(0.7)*	3.8	*(0.7)*	2.6	*(0.2)*	2.4	*(0.3)*
5–6-year girls	4.9	*(0.6)*	4.5	*(0.6)*	2.7	*(0.2)*	2.6	*(0.2)*
5–6-year boys	4.9	*(0.4)*	4.4	*(0.5)*	2.7	*(0.1)*	2.6	*(0.2)*
8–9-year girls	5.5	*(0.4)*	4.9	*(0.6)*	2.9	*(0.1)*	2.8	*(0.2)*
8–9-year boys	5.0	*(0.6)*	4.5	*(0.4)*	2.7	*(0.2)*	2.7	*(0.2)*
12–13-year girls	5.2	*(0.5)*	4.8	*(0.8)*	2.9	*(0.2)*	2.7	*(0.2)*
12–13-year boys	5.5	*(0.4)*	4.9	*(0.6)*	2.9	*(0.1)*	2.7	*(0.2)*
All children	4.9	*(0.7)*	4.4	*(0.8)*	2.7	*(0.2)*	2.6	*(0.3)*

[a]Consistency scores range from 1 (completely inconsistent) to 6 (repeated in all four interviews). Inconsistency scores range from 1 (multiple inconsistencies) to 3 (no inconsistencies, i.e. totally consistent).

in all four interviews), and the values of the *inconsistency* score ranged from 1 (multiple inconsistencies) to 3 (never inconsistent). Both the consistency and inconsistency scores were analysed by means of an ANOVA with Age (5 levels) and Gender (2 levels) the between-subjects factors and Event the within-subject factor (see Table 2).

For the consistency score, all three main effects were significant. Children became more consistent with age, $F(4,86) = 27.91$, $MS_e = 0.47$. The means for the age groups from youngest to oldest are as follows: 3.3, 4.3, 4.7, 5.0, and 5.1. It should be remembered that complete inconsistency is a score of 1.0, recall that is never contradictory albeit occasionally forgotten has a score of 4 or 5, and consistent recall in all four interviews is represented by a score of 6.0. Thus, even 2-year-olds are at least moderately consistent, whereas children who are 3 years of age or more are quite consistent in their recall of the same information on multiple occasions. Paired comparisons showed that each age group was significantly different from the others ($ps < 0.02$), with the exception of the oldest two groups which did not differ from each other. Girls ($M = 4.6$) are also more consistent in their recall than are boys ($M = 4.3$), $F(1,86) = 7.12$, $MS_e = 0.47$. As well, children recall their injury ($M = 4.9$) more consistently than they recall their hospital treatment ($M = 4.4$), $F(1,86) = 47.28$, $MS_e = 0.19$. There were no significant interactions.

Results of the analysis of mean inconsistency scores are parallel: all three main effects are significant, and none of the interactions are. Children are less inconsistent with age, $F(4,86) = 22.38$, $MS_e = 0.06$, with means (for the youngest to oldest children, respectively) as follows: 2.2, 2.5, 2.7, 2.8, and 2.8. Paired comparisons showed that each age group was significantly different from the others ($ps < 0.02$), with the exception of the oldest two groups which did not differ from each other. As well, girls ($M = 2.7$), are less inconsistent than are boys ($M = 2.5$), $F(1,86) = 8.23$, $MS_e = 0.06$, and children are less inconsistent when recalling the details of their injury ($M = 2.7$) than their hospital treatment ($M = 2.6$), $F(1,86) = 13.36$, $MS_e = 0.03$.

Table 3. The number of items that were consistently recalled by the children in two, three, or all four interviews that were correct versus errors, and percentage correct

	Number of interviews in which the same information is recalled								
	4 interviews			3 interviews			2 interviews		
Group	Correct	Error	%	Correct	Error	%	Correct	Error	%
2-year olds	44	0	100.0	33	2	94.3	27	5	84.4
3-year olds	204	1	99.5	60	5	92.3	33	6	84.6
5-year olds	361	1	99.7	100	4	96.2	67	10	87.0
8-year-olds	348	0	100.0	66	6	91.7	47	7	87.0
12-year olds	295	1	99.7	71	2	97.3	33	1	97.1
Total	1252	3	99.8	330	19	94.6	207	29	87.7

However, except for 2-year-olds, note that the older children are near the ceiling score of 3.0. That is, although the previous analysis of consistency scores showed considerable improvement with age, the analysis of inconsistency scores suggests that the children seldom explicitly contradicted themselves but rather were more likely to omit details in some interviews.

Accuracy of consistently recalled items

How accurate is information that is repeated over and over in multiple interviews? After all, children could be consistent and still be wrong. Such a pattern was sometimes found by Schwarzmueller (1997, unpublished manuscript), who found that children sometimes continued to repeat erroneous information in a subsequent interview. Table 3 shows the percentage accuracy of the details that children consistently reported in all four interviews, as well as the accuracy of those they sometimes omitted but when they were recalled, the same information was reported, i.e. they were consistent in the two or three interviews in which these details were recalled.

Details that are consistently recalled in every interview are virtually always accurate. Out of 1255 details that were consistently reported by the children in all four interviews, only three were commission errors that were repeated over time. One 4-year-old consistently claimed that she had cried a lot whereas her mother said that although she had cried, it was not really a lot; one 5-year-old claimed that there had been lots of blood whereas his mother said that although it had bled, it was not very much; and one 13-year-old claimed that his injury had been caused by being tripped by another child whereas the witness stated that the child had tripped over his own feet and that another child had not caused the accident. Interestingly, details that were less salient to the children and therefore sometimes omitted in their recalls were more likely to be wrong, although the accuracy of such details was still quite high. If the children recalled something three out of four times they were interviewed but did not report it during one interview, approximately 95% of the details they consistently recalled were accurate. If, on the other hand, they recalled it the same way in two interviews but did not report it during two other interviews, about 88% of the details were accurate. Thus, as details become less salient and less frequently recalled, there is more likelihood of error. However, details that are salient enough to be consistently recalled in every interview are almost always correct.

In summary, the children are impressively consistent in their reports of their injuries and hospital treatment. Specifically, the vast majority of the information that they recall at each interview is the same information that they recalled on previous occasions. Although older children provide a higher proportion of repeated information in their interviews, it is nevertheless the case that all ages of children are mostly repeating the same information over time in their widely separated interviews, especially about the injury event. Even though the last two interviews are separated by a year and the last interview took place fully two years after their injuries had occurred, the children are mostly providing the same information to the same questions. When average consistency scores are assigned to the children, the scores are high for all age groups except possibly the 2-year-olds. Furthermore, if something is salient enough to be recalled in every interview, it is almost invariably accurate whereas the less frequently it is recalled (or the more frequently the detail is forgotten), the higher the possibility that the detail is wrong, even if the same information had been consistently given in those interviews where the child did recall it.

Accuracy of 'new' information

The average number of 'old' and 'new' details in each interview is presented in Table 4. Children recalled almost twice as many 'old' details about the injury event as about the hospital event in each interview. Specifically, there are on average about 14, 15, and 16 details recalled by each child about the injury event in the 6-month, 1-year, and 2-year interviews, respectively. This compares with only approximately 7, $8\frac{1}{2}$, and 9 details about the hospital event in each interview. Nevertheless, children produce approximately the same number of new details about both events in each interview, namely 1 to 2 new details. Thus, there are very few new details provided by the children in any interview. This is especially true for the 12–13-year-olds, who produce virtually no new information, especially about their injury, in their last two interviews.

The proportion of 'new' information that was accurate was compared with the proportion of 'old' information that was accurate, to see if children's accuracy rates differ depending upon whether a detail was recalled in a previous interview or not. With so few new details about each event added by the children (and so many children providing no new information at all), it was not practical to calculate an ANOVA on the proportion of new information that is accurate. Thus, the data from all of the children in each age group were summed and a series of χ^2s were calculated on the number of correct and incorrect 'new' details. This was done for each event separately. Such a χ^2 calculation could be done for the 2-year-olds, 3–4-year-olds, and 5–6-year-olds, but the oldest two age groups provided so little new information that it was necessary to combine these two age groups for the χ^2s if we wished to analyse the two events separately.

Table 4 shows whether the percentage accuracy of 'old' information in an interview differs from the percentage accuracy of 'new' information, when 'new' versus 'old' details are compared within each event. Thus, for example, at the 6-month interview all age groups have higher accuracy proportions when providing 'old' information about the injury event than when providing 'new' information about the same event. However, during the same interview, there was no difference in the accuracy proportions of 'old' versus 'new' information about the hospital event (except when the data from all children are summed). During the 1-year interview, 'old' information about the injury was always more accurate than 'new' information about the injury, but in contrast to the 6-month interview, the children's 'old' information about the hospital event was also more accurate

Table 4. The mean number of 'old' and 'new' details about injury and hospital events provided by children in interviews at 6 months, 1 year, and 2 years, as well as the percentage of these details that are correct

	Injury				Hospital			
	'Old' information		'New' information		'Old' information		'New' information	
Age	*M*	%correct	*M*	%correct	*M*	%correct	*M*	%correct
6-month interview								
2 years	7.2	87%**	3.5	46%	3.7	73%	1.9	71%
3 years	13.6	89%**	3.6	75%	6.8	88%	2.3	76%
5 years	15.1	93%**	3.1	81%	7.5	89%	2.8	85%
8 years[a]	16.2	91%**	1.0	36%	7.6	93%	2.2	86%
12 years[a]	13.9	95%**	0.8	87%	9.3	89%	1.3	87%
All ages[b]	13.9	92%**	2.3	69%	7.3	89%*	2.2	82%
1-year interview								
2 years	10.3	80%**	2.3	28%	4.7	67%*	2.0	41%
3 years	15.3	89%**	2.1	58%	8.1	83%**	1.6	55%
5 years	15.9	91%**	2.1	63%	9.6	87%**	1.3	69%
8 years[a]	16.4	94%**	0.7	37%	8.9	89%	0.4	67%
12 years[a]	14.7	93%**	0.2	100%	10.2	87%	1.0	85%
All ages[b]	15.0	91%**	1.4	53%	8.7	86%**	1.2	63%
2-year interview								
2 years	10.6	79%**	3.4	50%	5.1	64%	1.3	36%
3–4 years	16.5	91%**	2.7	45%	8.9	80%**	2.0	53%
5–6 years	17.9	90%**	2.2	41%	10.3	85%**	1.8	58%
8–9 years[a]	17.4	92%**	1.0	38%	8.7	85%	0.7	73%
12–13 years[a]	13.9	95%**	0.1	50%	9.8	86%	0.5	89%
All ages[b]	15.9	91%	1.8	44%	9.0	83%**	1.3	58%

[a]The data from the 8–9 and 12–13-year-olds were combined for the χ^2s that compared the accuracy of 'new' and 'old' information about each event separately.
[b]The data are the mean for all the children pooled together into one group, not the arithmetic mean of the scores of the different age groups.
*'Old' information is more accurate than 'new' information provided about the same event at the $p < 0.05$ level.
**'Old' information is more accurate than 'new' information provided about the same event at the $p < 0.01$ level.

than their 'new' information, at least for the 2, 3–4 and 5–6-year-olds. The combined age groups of 8–12-year-olds were equivalently accurate for both 'old' and 'new' information. For the 2-year interview, 'old' information was always more accurate than 'new' information when the children recalled the injury event; however, when recalling the hospital event, only the 3–4- and 5–6-year-olds were more accurate when providing 'old' information (although the summed data from all children showed differences in accuracy).

In the current study, there was a very clear separation between two events: the injury event and the hospital event. Such a separate analysis of events is often not relevant in real-life situations such as forensic ones. In such instances, people want to know whether or not a new piece of information that was not recalled in previous interviews but is added in a later interview is likely to be accurate. To explore this, the new information provided by each age group during every interview was summed across events. Thus, all new details provided during each interview are included, regardless of the event being recalled. A series of goodness of fit χ^2s were calculated that compared the number of both accurate and incorrect new details with the numbers that would have been obtained had the responses been equivalently distributed between correct and incorrect. These χ^2s allowed us to determine whether the children's accuracy rates differed from chance responding. For the 2-year-olds, the accuracy of new information never differed from chance in any interview. That is, any new piece of information in any interview was just as likely to be wrong as right. For the 3–4-year-olds, new information was more accurate than chance at the 6-month interview, $\chi^2(df=1)=14.71$, $p<0.01$, but did not differ from chance accuracy rates during the later interviews at 1 and 2 years. For the 5–6-year-olds, new information was more accurate than chance responding during both the 6-month and 1-year interviews, $\chi^2(df=1)=37.94$, $p<0.01$ and $\chi^2(df=1)=4.50$, $p<0.05$, respectively, but did not differ from chance responding during the 2-year interview. For the 8–9-year-olds, new information was more accurate than chance at the 6-month interview, $\chi^2(df=1)=6.59$, $p<0.02$, but did not differ from chance accuracy rates during the later interviews at 1 and 2 years. For the 12–13-year-olds, new information was more accurate than chance during the 6-month interview, $\chi^2(df=1)=13.09$, $p<0.01$, but a goodness of fit χ^2 could not be calculated for the other interviews individually since there were only 3 and 2 incorrect new details provided by the entire group for the 1- and 2-year interviews, respectively. Consequently, their data were summed over the 1- and 2-year interviews. In these later interviews, new information was considerably more accurate than chance responding, $\chi^2(df=1)=9.68$, $p<0.01$. Thus, the oldest children were the only ones who consistently provided new information that was more accurate than chance responding, regardless of the delay between their injury and the interview.

Overall, when children talked about their injury, 'old' or previously recalled details were always more likely to be accurate than were 'new' details. However, this was not necessarily the case when children talked about their hospital treatment. Sometimes 'old' information was more accurate, and sometimes it was not—especially for both the youngest and oldest children. The 3–4- and 5–6-year-olds in particular have higher accuracy rates for 'old' information than 'new' information about the hospital, but only in their last two interviews. The oldest two groups never have differential accuracy rates when recalling the hospital event.

Furthermore, the 'new' information that the children add in their 6-month interview is more accurate than chance responding, except for the 2-year-olds for whom new details are always as likely to be wrong as right. However, at delays of greater than 6 months, the likelihood that a new detail is correct is equivalent to its likelihood of being wrong, with the exception of the 12–13-year-olds whose new details are always more accurate than chance. That is, for most age groups (with the exception of the 5–6- and 12–13-year-olds) the accuracy of new information did not differ from chance responding during the 1-year interview, and for no age group (except the 12–13-year-olds) did the accuracy of new information differ from chance at a delay of 2 years. Thus, new information added after delays that were at least a year was highly suspect from all ages of children up to 8–9 years of age.

DISCUSSION

Children in this study suffered an injury and were treated at a hospital ER, and over the next 2 years they were interviewed four times. What is striking is how consistent the children are in their reports of these complex events, and mostly 'old' or previously recalled information is present in every re-interview. This is especially true for those that took place long afterwards, namely a full year and even 2 years later. Even children as young as 2 years old at the time of injury are repeating much more 'old' information than generating new information in each interview. For 3–4-year-olds, fully three-quarters of their first re-interview, after a 6-month delay, is repeated information and this proportion rises considerably with older children. By the time children are 8–9 years old, the proportion of 'old' information rises to well over 90%.

If one assigns a consistency score (from 1 to 6) to every detail that children recall more than once, with scores between 1 and 3 reflecting various degrees of contradiction (1 representing maximal inconsistency) and scores between 4 and 6 reflecting complete agreement in the children's recall although varying on how many interviews contained target information (6 representing identical information in all four interviews), the children who were at least 3 years of age had mean consistency scores between 4 and 6. Only 2-year-olds had a mean consistency score that reflected contradiction (i.e. 3.3). Inconsistency scores (from 1 to 3) were assigned as well, and the children who were at least 5 years of age are near the ceiling score of 3 (representing no inconsistency at all), with 3–4-year-olds only slightly worse. Taken together, what the consistency and inconsistency scores reflect is that children's recall is mostly highly consistent and seldom contradictory. However, older children are more likely to include such consistent information in all interviews.

In summary, the children are mostly reporting the same details over and over in their interviews. These details are not simple agreements with yes/no questions; rather, the children themselves are generating content-rich information in response to wh- format questions. Inspection of the children's transcripts showed that children are particularly likely to consistently recall information about the core, central details of their experiences such as what happened when they got injured and the major things the doctor did to treat their injury. In contrast, they were more likely to sometimes omit or err in their recall of time details: time of day when the injury occurred, how long they cried, how long they waited in the ER waiting room, etc. They also had difficulty identifying people who played minor roles in their experiences.

The very high proportions of 'old' information in children's interviews contrast with the results of other studies. Since the same questions are asked in each interview in this study, it is not surprising that children's recall is much more consistent than in the interviews of Fivush and Hudson, among others (Fivush and Hamond, 1990; Fivush and Shukat, 1995; Hudson, 1990; Hudson and Fivush, 1991) in which questions often differed substantially between interviews. Less straightforward are differences with Steward *et al.* (1996) and Salmon and Pipe (1997), who used the same questions when re-interviewing. Steward *et al.* found that only half the details in each later interview had been mentioned previously and only a quarter was present in all three interviews. Salmon and Pipe found that the proportion of 'old' information was considerably less than half unless the children also had physical props available, and even these only increased their proportions to about half.

Why are children's reports so much more consistent in this study than in others? An important contributor may be the nature of the target events. First, the children's injury experiences were highly salient. Indeed, ratings by adult witnesses of the children's emotional reactions at the time of injury and treatment suggest that they were emotionally distressed, some hysterically so (Peterson and Bell, 1996). Such saliency and personal relevance have been shown to increase memory (Christianson, 1992). Furthermore, events were distinctive, the injury more than hospital treatment. Children's injury events were unique events in their lives, and although some aspects of hospital treatment were unique (such as having a bone cast or sutures), other aspects were not (waiting in the waiting room, being checked by the doctor, etc.). Distinctiveness and uniqueness are also important qualities of memorable experiences (Howe, 1997), and the greater distinctiveness of the injury event in comparison with the hospital event may account for why the injury is *recalled better* than the hospital event (Peterson, 1999; Peterson and Bell, 1996), as well as *recalled more consistently*, as found here. The decreased distinctiveness of Salmon and Pipe's target event, the check-up of a 'sick' teddy, may partially account for why their children were less consistent in recall. Indeed, those authors suggest that some of their inconsistency was due to children recalling related-but-different medical events. In the case of Steward *et al.*'s subjects, the target check-ups were by no means unique events. Since the children had other check-ups, including in the interval between the target check-up and the interview, Steward *et al.* provided an easily identifiable elephant logo for their clinic visits in an attempt to help children differentiate the target medical visit from others, an attempt that was not always successful.

There are other factors in Steward *et al.*'s research that may have decreased consistency in comparison with the current study. The children were re-interviewed by the same interviewer, and even preschoolers are sensitive to the knowledge state of conversational partners and are less likely to provide information that is already known (Fivush and Schwarzmueller, 1995; Mening-Peterson, 1975). Also, children experienced body touch that some were embarrassed about, and those who demonstrated shame were less likely to be consistent when recalling bodily touch.

Overall, this research suggests that it is important to investigate children's recall of a wide range of events. Even the two events investigated here, injury and hospital treatment, differed in how consistently children recalled relevant details and the differences between these two events and those investigated by other researchers are even more substantial. For some types of events, children may provide highly consistent reports even though long periods of time may have elapsed between interviews. Such high consistency is possible even for preschoolers.

An important forensic issue is not only consistency but also accuracy of consistently reported information. In this study children (at all ages) were virtually always accurate if they recalled the same information during all interviews. However, if a detail was less memorable (or less often reported), it was less likely to be accurate. If children recalled a relevant detail identically in three of the four interviews but did not recall it in one, the accuracy rate fell to 95%—still extremely high, but consistency was no longer a virtual guarantee of accuracy. If a detail was recalled identically in only two interviews and not reported in the others, accuracy fell to 88%. Thus, if something is routinely recalled, it seems to be reliably accurate. This difference in accuracy depending upon frequency of recall was also noted by Steward *et al.* (1996).

The appearance of new information in subsequent interviews is common, and it is important for researchers to give guidance to investigate and legal professionals about the reliability of such late-appearing information. The results of previous studies have suggested that such new information may be suspect because it is frequently inaccurate. In the present study, we found that new information in the 6-month interview was reliably more accurate than chance but that for most age groups, new information in the 1 or 2 year interviews was not. (A notable exception were the 12–13-year-olds, who provided very few new details, but accurate ones.)

In some ways the findings of this study are applicable to forensic situations, but there are difficulties with such a generalization. One potential problem is that in forensic situations there may be pressures applied to children to hide information, or they may feel shame or guilt and thus not report some events. In the present research, in contrast, children were very willing to talk about their experiences. The only component that might have engendered shame was their degree of crying, and we found that older boys in particular sometimes minimized or denied it. Another difficulty is that the interviews themselves may differ. In forensic situations, children are often not optimally interviewed and may be exposed to biased or misleading questions (Ceci and Bruck, 1993, 1995). Furthermore, a preponderance of questions directed toward young children are yes/no in format (Brainerd and Reyna, 1996; McGough and Warren, 1994), and consistency of response to these questions seems to be much less. Indeed, other research suggests that yes/no questions seem to be particularly problematic for preschoolers (Peterson and Biggs, 1997; Peterson *et al.*, 1999; Poole and White, 1993).

Caveats notwithstanding, there are also ways in which the findings of this study are relevant to forensic situations. There are frequently cases in which children report consistent details over and over in multiple interviews, and the results of this study suggest that such consistency seems to be an indicator of high reliability. In this report, such consistency in frequently reported details almost always was associated with accuracy. Furthermore, the greater the delay between target events and generation of new information in yet another interview, the greater the likelihood that the new information was inaccurate. Although new information supplied by children at the 6-month interview was reasonably likely to be correct, new information that first appeared 1 or 2 years after the target event was no more likely to be right than to be wrong. In forensic situations, of course, this low accuracy rate is totally unacceptable. On the other hand, this study encourages one to have confidence in a child's report when the same details are repeated in different interviews over time. Furthermore, even preschoolers are capable of generating consistent reports, even if there are long time delays between interviews.

ACKNOWLEDGEMENTS

Preparation of this article was primarily supported by Grant OGP0000513 from the Natural Sciences and Engineering Research Council of Canada. Additional funding came from the Memorial University Undergraduate Career Experience Program. We extend our thanks to the Janeway Hospital and their Emergency Room staff, and to all the recruiters, interviewers, transcribers and data analysers who participated. We also thank Michael Bruce-Lockhart for technical assistance, and Mary Courage for her considerable help in initiating the study. And most of all we thank the parents and children who willingly allowed us into their homes and were so cooperative.

REFERENCES

Brainerd CJ, Ornstein PA. 1991. Children's memory for witnessed events: The developmental backdrop. In *The Suggestibility of Children's Recollections: Implications for eyewitness testimony*, Doris J (ed.). American Psychological Association: Washington, DC; 10–20.

Brainerd CJ, Reyna VF. 1996. Mere memory testing creates false memories in children. *Developmental Psychology* **32**: 467–478.

Ceci SJ, Bruck M. 1993. The suggestibility of the child witness: A historical review and synthesis. *Psychological Bulletin* **113**: 403–439.

Ceci SJ, Bruck M. 1995. *Jeopardy in the Courtroom.* American Psychological Association: Washington, DC.

Christianson S-A. 1992. Emotional stress and eyewitness memory: A critical review. *Psychological Bulletin* **112**: 284–309.

Dent H. 1991. Experimental studies of interviewing child witnesses. In *The Suggestibility of Children's Recollections*, Doris J (ed.). Springer-Verlag: New York; 138–146.

Fivush R, Hamond NR. 1990. Autobiographical memory across the preschool years: Toward reconceptualizing childhood amnesia. In *Knowing and Remembering in Young Children*, Fivush R, Hudson JA (eds). Cambridge University Press: New York; 223–248.

Fivush R, Peterson C, Schwarzmueller A. in press. Questions and answers: The credibility of child witnesses in the context of specific questioning techniques. In *Memory and Suggestibility in the Forensic Interview*, Eisen ML, Goodman GS, Quas JA (eds). Erlbaum: Hillsdale, NJ.

Fivush R, Schwarzmueller A. 1995. Say it once again: Effects of repeated questions on children's event recall. *Journal of Traumatic Stress* **8**: 555–580.

Fivush R, Shukat JR. 1995. Content, consistency, and coherence of early autobiographical recall. In *Memory and Testimony in the Child Witness*, Zaragoza MS, Graham JR, Hall GCN, Hirschman R, Ben-Porath YS (eds). Sage: Thousand Oaks, CA; 5–23.

Goodman GS, Clarke-Stewart A. 1991. Suggestibility in children's testimony: Implications for sexual abuse investigations. In *The Suggestibility of Children's Recollections*, Doris J (ed.). American Psychological Association: Washington, DC; 92–105.

Home ML. 1997. Children's memory for traumatic experiences. *Learning and Individual Differences* **9**: 153–174.

Howe ML, O'Sullivan JT, Marche TA. 1992. Toward a theory of the development of long term retention. In *Development of Long-term Retention*, Howe ML, Brainerd CJ, Reyna VF (eds). Springer-Verlag: New York; 245–255.

Hudson JA. 1990. The emergence of autobiographical memory in mother–child conversation. In *Knowing and Remembering in Young Children*, Fivush R, Hudson JA (eds). Cambridge University Press: Cambridge; 166–196.

Hudson JA, Fivush R. 1991. As time goes by: Sixth graders remember a kindergarten experience. *Applied Cognitive Psychology* **5**: 347–360.

Hudson J, Nelson D. 1986. Repeated encounters of a similar kind: Effects of familiarity on children's autobiographic memory. *Cognitive Development* **1**: 253–271.

Leippe MR, Mannion AP, Romancyzk A. 1991. Eyewitness memory for a touching experience: Accuracy differences between child and adult witnesses. *Journal of Applied Psychology* **76**: 367–379.

McGough LS, Warren AR. 1994. The all-important investigative interview. *Juvenile and Family Court Journal* **45**: 13–29.

Menig-Peterson C. 1975. The modification of communicative behavior in preschool-aged children as a function of the listener's perspective. *Child Development* **46**: 1015–1018.

Peterson C. 1996. The preschool child witness: Errors in accounts of traumatic injury. *Canadian Journal of Behavioural Science* **28**: 36–42.

Peterson C. 1999. Children's memory for medical emergencies: Two years later. *Developmental Psychology* **35**: 1493–1506.

Peterson C, Bell M. 1996. Children's memory for traumatic injury. *Child Development* **67**: 3045–3070.

Peterson C, Biggs M. 1997. Interviewing children about trauma: Problems with 'specific' questions. *Journal of Traumatic Stress* **10**: 279–290.

Peterson C, Dowden C, Tobin J. 1999. Interviewing preschoolers: Comparisons of yes/no and wh-questions. *Journal of Law & Human Behavior* **23**: 539–555.

Poole DA. 1995. Strolling fuzzy-trace theory through eyewitness testimony (or vice versa). *Learning and Individual Differences* **17**: 87–93.

Poole DA, White LT. 1991. Effects of question repetition on the eyewitness testimony of children and adults. *Developmental Psychology* **27**: 975–986.

Poole DA, White LT. 1993. Two years later: Effects of question repetition and retention interval on the eyewitness testimony of children and adults. *Developmental Psychology* **29**: 844–853.

Poole DA, White LT. 1995. Tell me again and again: Stability and change in the repeated testimonies of children and adults. In *Memory and Testimony in the Child Witness*, Zaragoza MS, Graham JR, Hall GCN, Hirschman R, Ben-Porath YS (eds). Sage: Thousand Oaks, CA; 24–43.
Ross DF, Miller BS, Moran PB. 1987. The child in the eyes of the jury: Assessing mock jurors' perceptions of the child witness. In *Children's Eyewitness Memory*, Ceci SJ, Toglia MP, Ross DF (eds). Springer-Verlag: New York; 142–154.
Salmon K, Pipe M-E. 1997. Props and children's event reports: The impact of a 1-year delay. *Journal of Experimental Child Psychology* **65**: 261–292.
Steward MS, Steward DS, Farquhar L, Myers JEB, Reinhart M, Welker J, Joye N, Driskill J, Morgan J. 1996. Interviewing young children about body touch and handling. *Monographs of the Society for Research in Child Development* **57** (4–5, Serial No. 248).

