Journal of Experimental Child Psychology 109 (2011) 275–293

Contents lists available at ScienceDirect

Journal of Experimental Child Psychology

journal homepage: www.elsevier.com/locate/jecp

Children's memory reports over time: Getting both better and worse

Carole Peterson *

Department of Psychology, Memorial University of Newfoundland, St. John's, Newfoundland, Canada A1B 3X9

ARTICLE INFO

Article history:
Received 22 July 2010
Revised 19 January 2011
Available online 1 March 2011

Keywords:
Child memory
Eyewitness memory
Memory
Accuracy
Narrative detail
Reminiscence
Hypermnesia

ABSTRACT

Injured children (N = 145 between 2 and 13 years of age) were recruited from a hospital emergency room and were interviewed about the injury event soon afterward and then twice more at yearly intervals. Their transcripts were coded three ways: completeness of overall structural components of a prototypical injury event (e.g., who, when, where), amount of narrative detail (specifically unique units of information), and the accuracy of both types of information. Completeness components were also categorized as central or peripheral, and narrative details were coded as pertaining to persons, objects, attributes, locations, or activities. Over time, children maintained consistent completeness scores; that is, the overall structure of the event stayed the same. However, they provided more elaborative detail of all types and especially about attributes and activities. Only accuracy (of both types of information) deteriorated. Thus, different aspects of their interviews changed in different ways over 2 years. Implications for assessing changes over time in child witness reports are discussed.

© 2011 Elsevier Inc. All rights reserved.

Introduction

Children's long-term memory for highly salient and personally relevant events has been the focus of a number of studies (see reviews in Bauer, 2007; Peterson, 2002). Various studies have shown that children's memory reports get worse over time (Goodman, Hirschman, Hepps, & Rudy, 1991; Peterson, 1999; Peterson, & Whalen, 2001; Quas et al., 1999; Shrimpton, Oates, & Hayes, 1998), whereas others have found that they get better over time (Fivush, Sales, Goldberg, Bahrick, & Parker, 2004; Sales,

* Fax: +1 709 737 2430.
E-mail address: carole@mun.ca

0022-0965/$ - see front matter © 2011 Elsevier Inc. All rights reserved.
doi:10.1016/j.jecp.2011.01.009

Fivush, Parker, & Bahrick, 2005), and yet others have found that these reports remain equivalent across time (Baker-Ward, Gordon, Ornstein, Larus, & Clubb, 1993; Burgwyn-Bailes, Baker-Ward, Gordon, & Ornstein, 2001; Merritt, Ornstein, & Spicker, 1994; Salmon, Price, & Pereira, 2002). Such divergent conclusions are puzzling, and there have been few explicit attempts to clarify why such discrepancies may occur even though there is considerable forensic importance attached to understanding how children's memory reports change over time.

The premise motivating the current research is that at least some of these discrepancies may be related to methodological differences in how children's reports are assessed, specifically in how they are scored. Various investigators have coded different aspects of children's memory reports, and it may well be that these different aspects change in different ways over time. Unfortunately, to our knowledge, there has been little research that has looked at the same data from the same children using both the scoring procedures that have been associated with better memory over time and those that have been associated with worse memory over time. In the current study, we do just that. If the divergent conclusions of various studies are indeed partly related to how memory reports are coded, we can develop expectations about how children's memory reports may change over time – in what sorts of ways they may improve and, likewise, in what sorts of ways they may deteriorate. These more focused expectations may, in turn, assist those who need to predict or interpret children's testimony across time.

Change in memory reports over time has long been investigated (see recent reviews in Bauer, 2006, 2007), and decrements in amount and accuracy of recall are typically reported, although caveats abound. For example, less (or no) forgetting has been associated with shorter delays before recall, greater event salience and distinctiveness, information that involves core aspects or gist rather than peripheral details, and so on. Change over time is also relevant to theoretical concepts of reminiscence and hypermnesia. Reminiscence is the elicitation of new information in subsequent interviews, whereas hypermnesia is the increase in recall across multiple interviews (Bluck, Levine, & Laulhere, 1999; Bornstein, Liebel, & Scarberry, 1998; La Rooy, Pipe, & Murray, 2005). Hypermnesia is related to reminiscence in that hypermnesia requires new information to be provided, but at the same time there must also be less omission of previously recalled information than addition of new information such that the total amount of information increases across time. Investigators have routinely found that there is more information in memory than is elicited in any single recall attempt and, thus, the phenomenon of reminiscence is often seen (Gilbert & Fisher, 2006; La Rooy et al., 2005). However, whether or not there is a net increase of information (i.e., hypermnesia) is not reliable given that it is sometimes found and sometimes not found (see review in La Rooy et al., 2005).

Memory scoring

The three common methods of assessing children's recall of highly salient personal events involve (a) assessments of overall completeness, often via a checklist of relevant event features, (b) counting the number of unique units of information or narrative detail, and (c) the accuracy of recall, which would differ depending on what was coded in the study and, thus, could apply to either the accuracy of completeness components or the accuracy of narrative detail.

Recall completeness

This measure captures memory for the event as a whole. It assesses how many of the component aspects of the overall target event are recalled. In some studies, investigators compare children's recall with a checklist of component aspects to assess the number of event features that children provide about standardized medical procedures such as checkups and urinary catheterizations (e.g., Baker-Ward et al., 1993; Burgwyn-Bailes et al., 2001; Goodman et al., 1991; Merritt et al., 1994; Shrimpton et al., 1998). In addition, if one looks at how many components from the entire checklist are recalled, one can calculate the proportion of features recalled (i.e., completeness). A different but related way of coding memory reports was done by Peterson and her colleagues in studies on memory for injuries requiring hospital emergency room treatment (Peterson, 1999; Peterson & Bell, 1996; Peterson & Whalen, 2001). In contrast to the studies mentioned above, the authors did not have a standardized event with on-the-spot records of what had happened. Instead, they assessed recall completeness by developing a standardized prototype of an injury event and a different prototype of a typical hospital treatment event. Through interviewing adult eyewitnesses, they could determine which of the prototype components applied to any particular child's experience (which varied between children), and using this list, they then assessed the proportion of prototype components that children provided for each type of event. Thus, they could derive a measure of overall recall completeness.

When the completeness of children's reports is assessed, there tends to be relatively little change over time, although some investigators have found small decreases. For example, Baker-Ward et al. (1993) and Merritt et al. (1994) found no decrease in the number of features of medical procedures that were reported by 3- to 7-year-olds after a 6-week delay, and Burgwyn-Bailes et al. (2001) found no significant forgetting after a 1-year delay. Likewise, Peterson and her colleagues (Peterson, 1999; Peterson & Whalen, 2001) found no decrease in the percentage of components of a prototypical injury event that was recalled after either 2 or 5 years.

Narrative detail

This measure tabulates the number of unique units of information that children provide about people, activities, objects, descriptors, and locations. For example, this has been applied to children's recall of highly salient medical events such as urinary catheterization (Goodman, Quas, Batterman-Faunce, Riddlesberger, & Kuhn, 1994, 1997; Quas, Bauer, & Boyce, 2004; Quas et al., 1999). A variation is the number of subject–verb propositions that provide new information, and this has been used to evaluate memory reports for highly salient events such as a destructive hurricane (Bahrick, Parker, Fivush, & Levitt, 1998; Fivush et al., 2004; Sales et al., 2005). In this method of scoring memory reports, the investigator does not have a predeveloped list of features or prototype components to be recalled; rather, the amount of narrative detail in the report is tabulated. This gives one a sense of how extensive or embellished children's reports are. When investigators use this method of scoring, sometimes substantial improvements in children's memory over time have been reported. For example, 6 years after Hurricane Andrew, children provided considerably more information about the event than they had originally provided (Fivush et al., 2004; Sales et al., 2005). These children demonstrated both reminiscence (i.e., the recall of new information) and hypermnesia (i.e., the total amount of information increased).

Comparing recall completeness and narrative detail scoring systems

Completeness of recall and number of narrative details (or unique units of information) have a different focus, although they are complementary ways of assessing children's recall. When scoring recall completeness, each prototype component is scored as present or not regardless of how much narrative detail is provided. However, when scoring for narrative detail, each new detail is counted separately. Compare "We were at my Nan's" with "We were at my Nan's behind the green shed that's next to her house." These are equivalent under the prototype completeness scoring given that both state the location of described events and location is one of the component features of the prototype, but the latter contains considerably more narrative detail. Both ways of assessing children's recall are important forensically, even though they provide a different perspective. When children provide a relatively complete account of a target experience, they include information about who, what, when, where, and the sequence of events that occurred. In other words, they provide information about the overall structure of the event, which is considerably more helpful in forensic situations than someone who relates only a few components of the event, even if those components are recalled accurately. However, children's accounts are even more useful if they are rich in descriptive detail rather than being sparse skeletal accounts. In addition, children who provide rich descriptive detail about events are more likely to be believed as witnesses (Bala, Lee, & McNamara, 2001).

Currently, little research directly compares these two ways of assessing memory. An exception is Baker-Ward, Ornstein, Gordon, Follmer, and Clubb (1995), who scored 5- and 7-year-olds' recall of a pediatric examination in ways that parallel both the completeness and narrative detail methods described above. For completeness, they assessed the proportions of specific features (defined by a predetermined checklist) of a well-child examination that was provided by children, and for amount of information, they counted the number of units of information that were provided by children (scored similarly to the current study). Importantly, their comparison of the two coding systems was limited to data from only part of the interview. Specifically, they analyzed only children's responses to initial free recall questions and omitted responses to all of the other questions asked in the interviews. The investigators found that children's age was a significant factor in both types of analyses and that little incorrect information was provided. They also compared the rank orderings of children when scored with both coding systems, and the correlation was quite high. Their conclusion was that general findings about children's memory, such as age effects and overall memory competence, were unlikely to be artifacts of procedural differences in scoring. In addition, they emphasized that the two scoring systems targeted different aspects of child recall and, thus, provide differing, albeit complementary, views of child memory. However, they did not trace children's performance across time, and the age range for children was limited.

Accuracy

Accuracy has typically been assessed by tabulating commission errors and calculating either the number of such errors or the proportion of relevant information provided by children that is correct (i.e., not commission errors). The number of commission errors (as well as the proportion of correct information) has been assessed when both completeness scoring procedures and narrative detail procedures are used. In those studies where there was a checklist or video of medical or experimental procedures, children's memory was compared with records collected at the time. For interviews about unexpected and idiosyncratic injuries, children's reports were compared with accounts from adult witnesses. Another method that has been used is to ask parents to review transcripts of child interviews. Unlike findings with either completeness or amount of narrative detail, changes in children's accuracy (specifically the number of commission errors) have consistently been reported to reflect accuracy decrements across time because children make increasing numbers of commission errors in their recall. However, just as there have so far been no direct comparisons of children's memory reports scored using both completeness and narrative detail procedures, there also have been no comparisons of accuracy rates when what is being scored is so different. That is, scoring a global measure of "location" for accuracy under the completeness scoring system (i.e., whether or not a location is provided) is different from scoring for accuracy all the narrative descriptive details that may also be provided about that location. For example, a child may identify her injury as having taken place at her Nan's, and under the completeness scoring system this is scored as accurate because it happened there. However, even though the child correctly identified the location as Nan's, she may well have incorrectly described the specific place where events took place (e.g., by identifying Nan's front yard by her fence when events actually happened in the back yard that has no fence). These commission errors are captured by the narrative detail scoring system but not by the prototype component completeness system. The current study uses both the completeness and narrative detail scoring methods and assesses accuracy according to both systems of scoring by tabulating commission errors as defined by each system and then calculating the proportion of recall that is correct. This will provide a more differentiated assessment of how children's memory reports change across time.

Event being recalled

One of the most consistent findings in the literature is that the event being recalled affects the quality of memory reports. In a review of children's recall of autobiographical events, Peterson (2002) found that event salience and emotionality were robust predictors of the quality of long-term recall. As contrasting examples, preschoolers recalled approximately 20–30% of the features of a living room camping trip when interviewed only 1 day or 3 weeks later (Boland, Haden, & Ornstein, 2003; Haden, Ornstein, Eckerman, & Didow, 2001; Ornstein, Haden, & Hedrick, 2004), but they recalled 75% of the features of facial surgery events (caused by injuries) when they were interviewed one full year later (Burgwyn-Bailes et al., 2001). Distinctiveness of events also is important. Events that are similar to others that have frequently happened are less likely to be recalled, as are the details of individual repeated events. In terms of the current study, a unique and highly salient event is the target event being recalled – an event that also elicited considerable emotion at the time it occurred (Peterson, 1999; Peterson & Bell, 1996; Peterson & Whalen, 2001). Specifically, the target event is an injury serious enough to require hospital emergency room treatment. Although some prior studies have also looked at children's recall of their hospital treatment, recall of this event is not included here because the hospital event contains a lot of components that are highly similar to components of other visits to that emergency room for other reasons such as visits due to illness. Indeed, children of comparable ages have visited this emergency room an average of 12 times prior to the visit necessitated by the injury that was the target event (Peterson & Bell, 1996).

Hypotheses

The primary hypothesis is that different aspects of children's memory reports will change differentially over time. First, accuracy will decrease, regardless of whether accuracy of prototype components or narrative details is measured, in keeping with findings from a host of prior studies. This will be reflected by increasing proportions of recalled information that are commission errors. Second, completeness will remain fairly constant over time, as was found in previous studies about similar events (Peterson, 1999; Peterson & Whalen, 2001). That is, children will continue to provide a similar amount of injury prototype components, reflecting similarity in how well the overall event is recalled. Third, in contrast, the amount of narrative detail provided by children will increase over time. We predict that children's reports will reflect reminiscence in that we doubt any single report will provide all of the details of children's memories of these complex events. We also predict hypermnesia, parallel to findings about children's reports about hurricane experiences (Fivush et al., 2004; Sales et al., 2005). In addition, we will also investigate patterns of change for different types of narrative detail. We have no theoretically driven predictions about differences between various types of detail, but in applied settings such as forensic ones, it may be useful to know what sorts of details children are especially likely to recall more of over time or whether some sorts of detail are particularly likely to deteriorate in accuracy. This aspect of the study is strictly exploratory.

Method

Participants

Children were recruited from the emergency room of a children's hospital. Because medical care is free in Canada, the sample was a cross section of the community. Children had experienced trauma injuries, mostly bone fractures or lacerations requiring suturing. A total of 201 children participated in the initial and 1-year interviews: 21 2-year-olds (16 girls and 5 boys, mean age 2 years 6 months), 65 3- and 4-year-olds (29 girls and 36 boys, mean age 4 years 1 month), 68 5- and 6-year-olds (29 girls and 39 boys, mean age 5 years 10 months), 23 8- and 9-year-olds (12 girls and 11 boys, mean age 9 years 1 month), and 24 12- and 13-year-olds (9 girls and 15 boys, mean age 12 years 10 months). However, at the 2-year follow-up interview, only 145 children participated: 17 2-year-olds (12 girls and 5 boys, mean age 2 years 5 months), 39 3- and 4-year-olds (19 girls and 20 boys, mean age 4 years 0 months), 48 5- and 6-year-olds (23 girls and 25 boys, mean age 5 years 10 months), 21 8- and 9-year-olds (11 girls and 10 boys, mean age 8 years 11 months), and 20 12- and 13-year-olds (8 girls and 12 boys, mean age 12 years 7 months). Thus, only 145 children had all three interviews. Participant loss was due to inability to make contact with families (42 participants) or families choosing to no longer participate (14 participants). Of the 201 participants, the data of 96 children had been included in Peterson's (1999) study, specifically 11 2-year-olds, 18 3- and 4-year-olds, 26 5- and 6-year-olds, 22 8- and 9-year-olds, and 19 12- and 13-year-olds.

Procedure

Families were approached in the emergency room, and a research assistant described the study and gave them a written description to take home. Interested families provided contact information. A few days later, families were contacted to answer questions and set up home visits. During those visits, children and parental witnesses were interviewed independently using the same interview protocol. In the interview, free recall was elicited first ("Tell me about what happened when you/your child got hurt"). Subsequently, interviewees were asked a series of mostly "wh-" questions about the injury (e.g., where and when did it occur, what caused it) and then the hospital visit (e.g., what did the doctor do), although this article focuses only on the injury event because multiple aspects of the hospital visit were similar to many other prior hospital visits for illness and so on by children (Peterson & Bell, 1996). (See Appendix for topics of interview questions.) Both 1 and 2 years later, parents were recontacted for follow-up interviews, and during the phone conversation they were asked to refrain from talking to their children about target events prior to the researcher's visits to their homes. The same interview was administered to children at all three visits. All aspects of the study were approved by the university's human investigation committee for ethical treatment of human participants.

Data coding

Recall completeness

Idealized prototypes of typical injury events were developed in earlier research (Peterson & Bell, 1996), and interview questions queried all components of these prototypes. In addition, all prototype components were identified as central or peripheral according to Heuer and Reisberg's (1992) distinction of plot-relevant versus plot-irrelevant information. In Peterson and Bell (1996), peripheral details were also subdivided into two categories, but these are combined here. (See Appendix for classification of prototype components as central versus peripheral.) Although every child's injury experience was unique in some ways, it nevertheless conformed to this idealized prototypical pattern. Through searching parent transcripts, one could determine which prototype components applied to each child. Although most prototype components applied to all children (e.g., the injury occurred at a specific time and location), some prototype components applied to only a subset of the children (e.g., "objects of home treatment" was not relevant to a child who was immediately rushed to the hospital after injury rather than taken inside the home first for initial treatment). Because there were different numbers of prototype components that applied to each child, different children had different numbers of scorable components that could potentially be present in their recall. After determining which prototype components applied to each child, their transcripts were searched to determine whether the child supplied information relevant to each prototype component. If so, it was compared with information provided by parents to assess component accuracy. The completeness proportion was derived separately for central and peripheral components as well as combined to form a total completeness score. The completeness of a child's recall was calculated by dividing the number of component items correctly recalled by the number of component items that were relevant for that child according to parental report and, thus, could potentially have been recalled. For example, if parents identified 8 of the central components as relevant to their child's injury but their child correctly recalled only 6 of them, the child was given a proportion score of .75 for injury central completeness. Only components that were accurately recalled are presented here; errors in recalling prototype components are presented below. This scoring is identical to that used in prior studies about injuries requiring emergency room treatment (Peterson & Bell, 1996; Peterson & Whalen, 2001).

Accuracy of completeness components

Only commission errors were counted (i.e., instances in which a child stated information that was explicitly contradicted by parental report). For most children, only one parent was interviewed. For those few families in which both parents were interviewed, their information was never contradictory. Although using parental report as the "gold standard" for accuracy is not ideal, it is the only viable option for unexpected real-world events such as the ones used here. However, all parental information against which children's reports were compared was derived from the initial visit that took place a few days after injury. Parents were never reinterviewed; thus, potential deterioration of their reports over time is not an issue. Use of parent–child information convergence in this way to assess accuracy has been done in a number of previous studies (e.g., Peterson, 1999; Peterson & Bell, 1996; Peterson & Whalen, 2001). The numbers of errors about prototype components were separated into central versus peripheral information. Then the percentage accuracy of the components recalled by children was calculated (i.e., the proportion of those components that were accurate rather than commission errors). Instead of using the *possible* components that children potentially could have remembered as the denominator in calculations (as in the completeness measure described above), this measure used the *actual* components that children did provide as the denominator in calculations and then the proportion of those components that was accurate was derived. That is, the number of correct prototype components in each category was divided by the total number of components provided by children (i.e., the sum of correct ones plus commission errors). For example, if a child provided 6 correct central components of her injury and 1 commission error, her proportion accuracy score was .86. Scoring was identical to that used in previous studies (Peterson & Bell, 1996; Peterson & Whalen, 2001).

Unique narrative detail

Each unique unit of information introduced by the child was counted. The total number of unique units was tabulated and included details pertaining to person (e.g., "*Nan* picked me up"), location (e.g., "I was at the *park*"), action (e.g., "I was *running*"), object (e.g., "There was a *rock*"), and attribute (e.g., "It was *sharp*"). Scoring was the same as that developed by Fivush (1991) and used in a number of subsequent studies, including one about similar injury events (Peterson & Roberts, 2003). These unique information units were not divided into central versus peripheral details because of difficulty with classification. Consider the following: "I hit a piece of the porch that was sticking up and had nails in it because daddy had starting fixing it last weekend but he didn't finish it 'cause he ran out of nails and had to go to the store again." The object involved in the injury (piece of the porch) is a central component of the injury prototype, but is each of the elaborating details about that object, including the history of why it was in its state at that time, also perceived by children as central? Although Christianson and Loftus (1991) defined any detail or element that is *associated with* a central element as central rather than peripheral, others did not (Heuer & Reisberg, 1992) and there was no way of knowing whether a particular child did so.

Accuracy of unique narrative detail

Adult witness transcripts were searched to assess the accuracy of each unique unit of information provided by children (i.e., whether or not it was a commission error). Not all unique units could be assessed for accuracy; a mean of 87.5% (range = 47–100 for different children) of unique units could be coded as either accurate units or commission errors, and for the remainder adults did not provide relevant information that would allow coding for accuracy. Accuracy is defined as the number of details confirmed as correct divided by the number confirmed as correct plus commission errors.

Reliability of scoring

To establish reliability for completeness, prototype components were first identified in the children's transcripts. Two raters independently scored 15% of the transcripts, and agreement (scored as the number of agreements divided by the number of agreements plus disagreements) was 98% for the presence of scorable components in the transcripts (i.e., their completeness). Agreement for classification of components into central versus peripheral categories was 91%. To establish reliability for the accuracy of children's recall of prototype components, two raters compared 15% of the children's transcripts with parental transcripts to verify accuracy of identified components, and reliability (again the number of agreements over agreements plus disagreements) was 90%. One of the raters then scored the remainder of the transcripts for completeness and accuracy. For the number of unique units of information, two raters scored 15% of the transcripts; agreement was 92% for identification of unique units of information and 84% for accuracy of those units. One of the raters then scored the remainder.

Results

Data on children's recall of their injury experiences are presented separately for four types of data: (a) the completeness of their recall (i.e., the proportion of relevant prototype components that are recalled), (b) the accuracy of completeness components, (c) the number of narrative details, and (d) the accuracy of those narrative details. Data relevant to completeness are broken down into whether prototype components are central or peripheral. In addition, total completeness scores are presented; these are the sum of central and peripheral components and, therefore, present an overall picture of children's completeness. For analyses of narrative detail (both frequency and accuracy), the total amount of narrative detail (summing over the various types of information) is presented as well as data on the individual types of detail. Data are presented for all three interviews – initial, 1-year follow-up, and 2-year follow-up – and include all 201 children in the initial and 1-year interviews but only 145 children in the 2-year interview. Because more children participated in the initial and 1-year interviews, all analyses of variance (ANOVAs) were run twice: (a) on all 201 participants, with time having two levels (initial and 1-year interviews), and (b) on 145 participants, with time having three levels (all three interviews). In addition, the initial and 1-year scores of the 56 children who were not available after 2 years were compared with the scores of the 145 children who were, and these two groups did not differ. Because findings did not differ depending on whether 201 or 145 children were used, for simplicity, only analyses of the 145 children available for all three interviews are presented.

Completeness of recall

Table 1 shows the mean percentage completeness of children's recall during initial, 1-year, and 2-year interviews. Children's completeness scores were analyzed via a 5 (Age) × 2 (Gender) × 2 (Centrality) × 3 (Time) ANOVA with both age and gender as between-participants variables, centrality (central vs. peripheral) as a within-participant variable, and time as a repeated measure. Because there were no effects of gender alone or in interaction, data are not separated into gender in Table 1. Parallel to results reported previously with a smaller sample, children had more complete recall with age, $F(4, 135) = 37.70$, $p < .001$, $\eta^2 = .528$. The mean percentage recall of all prototype components for the different age groups was 51.0%, 67.6%, 78.8%, 81.2%, and 82.0% for the youngest through oldest groups, respectively. Post hoc Tukey HSD tests showed that 2-year-olds were less complete than 3- and 4-year-olds, and both of these groups were less complete than all three older groups, which did not differ from each other. In addition, information centrality was significant, $F(1, 135) = 92.63$, $p < .001$, $\eta^2 = .407$, with 78.5% of central components recalled but only 66.7% of peripheral components recalled. There were no other significant effects. Importantly, there was no significant deterioration in recall completeness over 2 years regardless of age.

Table 1
Recall completeness: Means (and standard deviations) of completeness component percentages (central, peripheral, and total components) during initial, 1-, and 2-year follow-up interviews

Age (years)	N^a	N^b	Centrality of information	Interview		
				Initial	1-Year	2-Year
2	21	17	Central	48.7 (16.7)	58.6 (22.8)	59.9 (22.6)
			Peripheral	40.7 (20.5)	50.0 (24.2)	47.4 (17.8)
			Total	44.6 (15.4)	54.6 (21.2)	53.8 (18.1)
3 & 4	65	39	Central	72.8 (18.0)	70.0 (20.9)	74.0 (16.3)
			Peripheral	57.4 (18.0)	63.1 (21.2)	65.4 (16.4)
			Total	66.3 (15.0)	66.1 (18.2)	70.5 (14.3)
5 & 6	68	48	Central	84.6 (10.7)	83.2 (16.5)	84.1 (15.0)
			Peripheral	72.7 (13.2)	74.8 (16.6)	70.2 (17.2)
			Total	79.2 (8.7)	79.3 (13.1)	77.9 (12.6)
8 & 9	23	21	Central	89.8 (11.5)	81.7 (15.2)	85.4 (11.0)
			Peripheral	79.8 (10.2)	70.2 (14.2)	76.1 (14.0)
			Total	85.5 (8.9)	76.8 (11.4)	81.2 (9.7)
12 & 13	24	20	Central	92.3 (10.1)	87.0 (9.6)	82.8 (12.9)
			Peripheral	77.4 (13.0)	73.8 (17.8)	75.2 (11.2)
			Total	85.6 (9.2)	81.3 (10.8)	79.0 (8.8)
All	201	145	Central	78.6 (18.7)	76.6 (19.9)	78.6 (17.5)
			Peripheral	65.8 (19.5)	67.8 (20.4)	67.9 (17.9)
			Total	72.9 (17.0)	72.4 (17.6)	73.7 (15.3)

Note. Data on 96 children from Peterson (1999) are included in the total information category in the initial and 2-year data, although that study did not break down the recalled information into central versus peripheral and did not report 1-year follow-up data.

[a] Number of participants in the initial and 1-year interviews.

[b] Number of participants in the 2-year interview.

Accuracy of completeness components

The accuracy of the components used to score recall completeness is shown in Table 2. Specifically, the proportions of components that were correct in comparison with the total number of components provided by children (i.e., both accurate ones and commission errors) are presented. These data were analyzed via a 5 (Age) × 2 (Gender) × 2 (Centrality) × 3 (Time) ANOVA with both age and gender as between-participants variables, centrality as a within-participant variable, and time as a repeated measure. Girls ($M = 88.9\%$) were significantly more accurate than boys ($M = 87.0\%$), $F(1, 134) = 4.97$, $p = .028$, $\eta^2 = .036$ (i.e., had proportionately fewer commission errors), although the absolute difference is small. Because gender interacted with no other factors and the difference was minimal, for simplicity, the data in Table 2 are not presented separately by gender. Older children were more accurate than younger children, $F(4, 134) = 16.90$, $p < .001$, $\eta^2 = .335$. The mean percentage accuracy of all prototype components for the different age groups was 74.9%, 81.9%, 85.2%, 86.5%, and 95.2% for the youngest through oldest groups, respectively. In contrast to the completeness analyses above, children became less accurate with time (i.e., made proportionately more commission errors), $F(2, 268) = 32.44$, $p < .001$, $\eta^2 = .195$. There was both a linear component, $F(1, 134) = 45.17$, $p < .001$, $\eta^2 = .252$, and a quadratic component, $F(1, 134) = 16.64$, $p < .001$, $\eta^2 = .110$, to the main effect of time, illustrating both that completeness accuracy decreased over time and that the sharpest decrease was between the first and second interviews. However, this pattern interacted with age given that there was a significant Time × Age interaction, $F(8, 268) = 2.33$, $p = .019$, $\eta^2 = .650$, although only the quadratic component was significant, $F(4, 134) = 3.18$, $p = .016$, $\eta^2 = .087$. This Time × Age interaction reflects the fact that the deterioration in accuracy (i.e., increase in commission errors over time) was different for children in different age groups and is depicted in Fig. 1.

The centrality of the prototype components also was significant. Children recalled central components ($M = 89.4\%$) better than peripheral ones ($M = 86.4\%$), $F(1, 134) = 9.86$, $p = .002$, $\eta^2 = .069$. However, this was complicated by a Centrality × Age interaction, $F(4, 134) = 3.25$, $p = .014$, $\eta^2 = .088$, depicted in Fig. 2. To explore this interaction, 2 (Centrality) × 2 (Gender) × 3 (Time) ANOVAs were calculated for each age group separately. Central components were likely to be recalled proportionately more accurately than peripheral ones for 2-year-olds ($p = .025$) and marginally more accurately for 3- and 4-year-olds ($p = .051$). Thus, only younger children were less accurate at recalling peripheral components; older children recalled both central and peripheral components with equivalent accuracy.

Table 2
Completeness accuracy: Means (and standard deviations) of accuracy percentages for completeness components (central, peripheral, and total components) during initial, 1-, and 2-year follow-up interviews

Age (years)	Centrality of information	Interview		
		Initial	1-Year	2-Year
2	Central	88.1 (10.1)	79.9 (22.8)	84.5 (10.7)
	Peripheral	79.3 (24.7)	63.6 (24.6)	69.8 (17.9)
	Total	84.3 (14.1)	73.6 (18.7)	74.9 (10.2)
3 & 4	Central	90.1 (11.0)	81.5 (17.3)	84.9 (14.4)
	Peripheral	85.8 (16.2)	79.2 (18.8)	78.1 (18.8)
	Total	87.9 (10.2)	80.6 (14.8)	81.9 (13.2)
5 & 6	Central	95.1 (6.9)	86.5 (12.7)	87.1 (13.8)
	Peripheral	93.1 (11.0)	88.1 (13.4)	82.7 (19.5)
	Total	94.1 (7.0)	87.2 (9.0)	85.2 (12.7)
8 & 9	Central	98.3 (5.1)	90.5 (9.9)	86.2 (12.0)
	Peripheral	97.9 (5.5)	92.1 (10.5)	87.0 (12.2)
	Total	97.9 (3.6)	90.9 (7.7)	86.5 (9.6)
12 & 13	Central	98.5 (5.4)	94.0 (8.3)	93.3 (10.0)
	Peripheral	97.8 (5.4)	93.3 (14.3)	97.4 (6.3)
	Total	98.2 (3.6)	93.9 (8.3)	95.2 (6.4)
All	Central	93.7 (9.1)	85.7 (15.2)	87.1 (13.1)
	Peripheral	90.7 (14.7)	84.3 (18.1)	83.2 (18.2)
	Total	92.1 (9.6)	85.0 (13.4)	84.9 (12.6)

Note. Data on 96 children from Peterson (1999) are included in the total category in the initial and 2-year data.

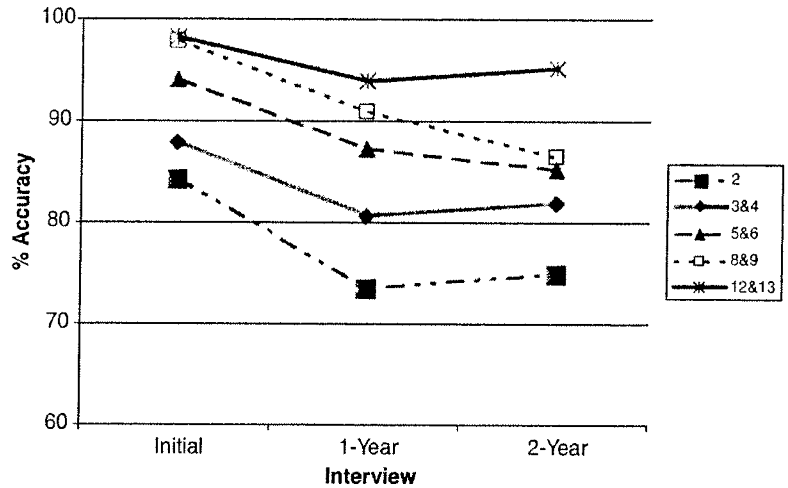

Fig. 1. Time × Age interaction for accuracy of completeness prototype components. Symbols show various age groups (in years).

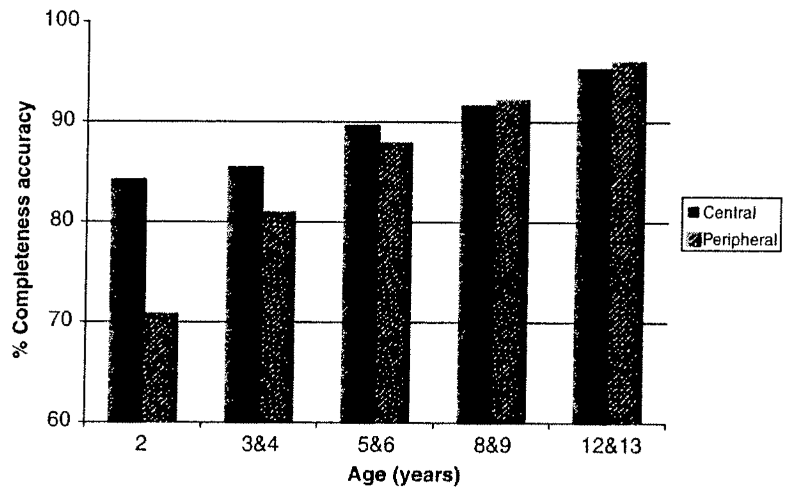

Fig. 2. Centrality × Age interaction for accuracy of completeness prototype components. Symbols show various age groups (in years).

Table 3
Total numbers of unique narrative details: Means (and standard deviations) for the number of narrative details provided during initial, 1-year, and 2-year follow-up interviews

Age (years)	Interview		
	Initial	1-Year	2-Year
2	18.6 (11.8)	19.9 (9.3)	25.5 (11.1)
3 & 4	33.2 (12.9)	38.9 (16.3)	45.3 (17.7)
5 & 6	43.8 (16.2)	42.3 (15.6)	57.2 (19.6)
8 & 9	50.5 (14.3)	49.8 (18.6)	63.4 (19.6)
12 & 13	54.9 (17.8)	51.3 (15.3)	58.2 (20.8)
All	40.8 (18.2)	41.4 (17.8)	51.7 (21.3)

Amount of unique narrative detail

The number of unique narrative details included by children in their memory reports has not heretofore been presented for children recalling emergency-room injuries, and to analyze how informative children's reports were, a 5 (Age) × 2 (Gender) × 3 (Time) ANOVA was computed on the total number of all unique details with both age and gender as between-participants variables and time as a repeated measure. These data are presented in Table 3. Because gender was not a significant factor alone or in interaction, data are not presented separately by gender. Older children were more informative than younger children (*Ms* = 21.3, 39.1, 47.8, 54.6, and 54.8 unique details for children in the youngest through oldest groups, respectively), $F(4, 135) = 20.26$, $p < .001$, $\eta^2 = .375$. Post hoc Tukey HSD tests showed that 2-year-olds provided fewer narrative details than 3- and 4-year-olds, and both of these groups provided fewer details than all three older groups, which did not differ from each other. There was also a significant effect of time, $F(2, 270) = 22.31$, $p < .001$, $\eta^2 = .142$. There was both a linear component, $F(1, 135) = 32.21$, $p < .001$, $\eta^2 = .193$, and a quadratic component, $F(1, 135) = 9.78$, $p = .002$, $\eta^2 = .068$, to the main effect of time, illustrating both that the number of unique narrative details increased and that there was only a small increase in the amount of narrative detail from the initial interview to the 1-year interview (40.8 vs. 41.8 details, respectively), but there was a large increase from the 1-year interview to the 2-year interview (51.7 details for the latter). Thus, with time, children provided more information, not less, about their injury experience.

The frequencies of the different types of narrative detail (persons, objects, attributes, locations, and actions) are presented in Table 4. To explore how the frequencies of the various types of narrative detail changed over time, a 5 (Age) × 2 (Gender) × 3 (Time) multivariate analysis of variance (MANOVA) was done on the five types of detail with time as a repeated measure. As with the ANOVA using the total number of narrative details (i.e., summing across the five different types of detail), there were significant main effects of both age and time. However, our interest here is with the different types of detail. The main effect of detail type was significant, Wilks' $F(4, 132) = 153.30$, $p < .001$, $\eta^2 = .823$, due to various types of detail having different frequencies. These frequencies ranged from 3.8 location details to 13.2 action details. In addition, the type of detail interacted with both time, Wilks' $F(8, 128) = 3.35$, $p = .002$, $\eta^2 = .173$, and age, Wilks' $F(16, 540) = 5.56$, $p < .001$, $\eta^2 = .169$.

To explore these interactions, separate 5 (Age) × 2 (Gender) × 3 (Time) ANOVAs were computed on each of the five different types of narrative detail. There was a main effect of time for all five types of detail: $F(2, 270) = 10.14$, $p < .001$, $\eta^2 = .070$ for persons, $F(2, 270) = 4.54$, $p = .011$, $\eta^2 = .033$ for objects, $F(2, 270) = 12.13$, $p < .001$, $\eta^2 = .082$ for attributes, $F(2, 270) = 14.76$, $p < .001$, $\eta^2 = .099$ for location, and $F(2, 270) = 11.68$, $p < .001$, $\eta^2 = .080$ for actions. However, although there was a linear component for all five types of detail – $F(1, 135) = 14.84$, $p < .001$, $\eta^2 = .099$ for persons, $F(1, 135) = 7.66$, $p = .006$, $\eta^2 = .054$ for objects, $F(1, 135) = 21.62$, $p < .001$, $\eta^2 = .138$ for attributes, $F(1, 135) = 18.08$, $p < .001$,

Table 4
Types of narrative details: Means (and standard deviations) for the number of narrative details provided about different content during initial, 1-, and 2-year follow-up interviews

Age (years)	Interview	Type of detail				
		Persons	Objects	Attributes	Locations	Activities
2	Initial	4.2 (2.3)	2.6 (2.6)	3.2 (4.1)	1.7 (1.2)	4.8 (3.2)
	1-Year	5.3 (2.1)	3.8 (3.2)	5.1 (4.1)	2.1 (1.5)	5.8 (3.5)
	2-Year	5.5 (3.1)	4.2 (2.2)	6.2 (3.9)	2.2 (1.4)	6.1 (3.8)
3 & 4	Initial	4.8 (2.6)	5.6 (3.4)	8.2 (5.9)	3.0 (1.7)	9.6 (5.0)
	1-Year	5.2 (2.1)	6.2 (2.9)	10.0 (6.8)	3.3 (1.5)	11.4 (5.2)
	2-Year	6.6 (3.0)	6.8 (3.0)	12.4 (6.7)	3.4 (1.4)	12.3 (5.9)
5 & 6	Initial	5.6 (2.3)	6.8 (3.4)	11.9 (6.1)	3.8 (1.6)	12.6 (6.1)
	1-Year	6.0 (2.6)	6.6 (3.2)	11.8 (6.4)	3.7 (2.0)	13.1 (5.7)
	2-Year	7.4 (3.7)	8.3 (4.4)	13.8 (6.0)	4.7 (1.7)	17.2 (6.8)
8 & 9	Initial	6.3 2.4)	7.6 (3.0)	9.5 (4.6)	4.3 (1.8)	16.1 (5.9)
	1-Year	6.0 (3.0)	7.1 (4.0)	10.5 (5.9)	4.3 (1.7)	15.3 (7.2)
	2-Year	8.1 (3.9)	8.3 (3.6)	13.2 (4.4)	6.2 (2.4)	19.7 (7.2)
12 & 13	Initial	6.8 (4.3)	8.6 (4.1)	14.2 (7.8)	4.8 (1.4)	19.3 (5.5)
	1-Year	6.3 (2.5)	7.5 (2.8)	13.2 (5.7)	4.8 (2.2)	17.5 (6.3)
	2-Year	5.5 (2.3)	8.4 (3.5)	13.4 (6.3)	5.6 (2.1)	18.2 (6.7)
All	Initial	5.4 (2.8)	6.3 (3.7)	9.8 (6.6)	3.5 (1.8)	12.0 (6.7)
	1-Year	5.7 (2.4)	6.3 (3.3)	10.5 (6.5)	3.6 (1.9)	12.6 (6.3)
	2-Year	6.8 (3.4)	7.4 (3.8)	12.4 (6.2)	4.4 (2.2)	15.1 (7.3)

$\eta^2 = .118$ for location, and $F(1, 135) = 13.75$, $p < .001$, $\eta^2 = .092$ for actions – there was also a quadratic component for location details, $F(1, 135) = 10.54$, $p = .001$, $\eta^2 = .072$, and actions, $F(1, 135) = 8.70$, $p = .004$, $\eta^2 = .061$. Thus, there was an increase over time for all types of detail, but for location and activity details, the increase was much greater between the 1-year and 2-year follow-up interviews than between the initial and 1-year follow-up interviews (see Table 4). In addition, the amount of increase varied for different types of detail; children provided approximately 1 new narrative detail of each type for people, objects, and locations but provided 3 new details about both attributes and actions.

In terms of age, there was not a main effect of age for all five types of detail; rather, there was a main effect of age only for objects, $F(4, 135) = 7.68$, $p < .001$, $\eta^2 = .185$, attributes, $F(4, 135) = 9.49$, $p < .001$, $\eta^2 = .219$, and locations, $F(4, 135) = 20.04$, $p < .001$, $\eta^2 = .373$. For both the number of object and attribute narrative details provided by children, Tukey HSD tests showed that 2-year-olds provided fewer details than all of the older groups, which did not differ from each other. For location details, with increasing age, each group provided more location details than children in younger groups with the exception of the two oldest groups, which provided the same number of details (*M*s = 2.0, 3.2, 4.1, 4.9, and 5.1 for the youngest through oldest groups, respectively). For ANOVAs on the number of object, attribute, and location details, there were no other significant effects besides time and age. In contrast, for the number of action details, there was no main effect of age but there was a significant Time × Age interaction, $F(8, 270) = 2.70$, $p = .007$, $\eta^2 = .074$. To explore this interaction, a 3 (Time) × 2 (Gender) ANOVA was calculated for each age group separately. For both the youngest and oldest groups, time was nonsignificant. However, there was a linear increase in the number of action details over time for 3- and 4-year-olds, and there was a quadratic increase over time for both 5- and 6-year-olds and 8- and 9-year-olds. There was also no main effect of age for the number of persons provided by children, although there was a significant Age × Gender interaction, $F(4, 135) = 3.32$, $p = .013$, $\eta^2 = .090$. To explore this interaction, the analysis was repeated for each gender separately. For girls, the number of narrative details about persons increased with age, $F(4, 68) = 4.38$, $p = .003$, $\eta^2 = .205$, but there was no significant age effect for boys ($p = .878$); rather, the number of person details remained the same across age for boys.

Accuracy of unique narrative detail

The percentage accuracy of children's narrative details is presented in Table 5. These data represent the proportions of narrative details that were correct in comparison with all recalled details (i.e., correct ones plus commission errors). Accuracy assessments could not be made on 12.5% of the unique narrative details provided by children. For those details that were able to be classified as correct or not, the proportion that was correct was analyzed in a 5 (Age) × 2 (Gender) × 3 (Time) ANOVA with time as a repeated measure. Because gender was not significant alone or in interaction, data on the accuracy of children's unique details are not presented separately by gender in Table 5. Children became proportionately more accurate (i.e., made proportionately fewer commission errors) with age, $F(4, 135) = 18.28$, $p < .001$, $\eta^2 = .351$. The mean percentage accuracy for unique narrative details was 77.5%, 84.5%, 89.8%, 94.4%, and 95.2% for the youngest through oldest groups, respectively. Post hoc Tukey HSD tests showed that 2-year-olds were proportionately less accurate than 3- and 4-year-olds, and both of these groups were proportionately less accurate than all three older groups, which did not differ from each other. The effect of time was also significant, $F(2, 270) = 17.37$, $p < .001$, $\eta^2 = .114$; only

Table 5
Accuracy of narrative details: Means (and standard deviations) for the percentage of details that were accurate during initial, 1-, and 2-year follow-up interviews

Age (years)	Interview		
	Initial	1-Year	2-Year
2	80.4 (19.8)	76.7 (12.4)	75.3 (13.2)
3 & 4	88.4 (10.3)	82.3 (15.2)	82.7 (12.2)
5 & 6	94.1 (6.3)	90.0 (7.3)	85.2 (10.0)
8 & 9	98.3 (2.8)	93.4 (8.8)	91.5 (9.8)
12 & 13	98.8 (3.5)	93.2 (6.9)	93.7 (5.4)
All	92.4 (10.7)	87.4 (12.0)	85.6 (11.6)

the linear (not the quadratic) component of time was significant, $F(1, 135) = 26.64$, $p < .001$, $\eta^2 = .165$. Accuracy of unique narrative details decreased from 92.4% to 85.6% over the 2 years of interviews.

The percentage accuracy of the different types of narrative detail (persons, objects, attributes, locations, and activities) is presented in Table 6. To explore how the accuracy of the various types of narrative detail changed over time, another Age × Gender × Time MANOVA was done on the five types of detail. As with the ANOVA using the total number of narrative details (i.e., summing across the five different types of detail), there were significant main effects of both age and time. Likewise, there was a significant effect of detail type, Wilks' $F(4, 132) = 36.96$, $p < .001$, $\eta^2 = .528$, due to various types of detail having different accuracy rates. These ranged from 88.8% for location details to 68.8% for attributes. In addition, the type of detail interacted with age, Wilks' $F(16, 404) = 2.43$, $p = .002$, $\eta^2 = .068$, and gender, Wilks' $F(4, 132) = 2.57$, $p = .041$, $\eta^2 = .072$. However, there was no interaction between detail type and time, suggesting that accuracy decreased equivalently for all types of narrative detail or, in other words, children were equally likely to increase the proportion of commission errors they made regardless of the type of content.

To explore the interactions between detail type and both age and gender, separate Age × Gender × Time ANOVAs were computed on each of the five different types of narrative detail. For each type of detail, the main effect of age was significant: $F(4, 135) = 8.15$, $p < .001$, $\eta^2 = .195$ for persons, $F(4, 135) = 8.12$, $p < .001$, $\eta^2 = .194$ for objects, $F(4, 135) = 14.99$, $p < .001$, $\eta^2 = .308$ for attributes, $F(4, 135) = 6.01$, $p < .001$, $\eta^2 = .151$ for location, and $F(4, 135) = 10.42$, $p < .001$, $\eta^2 = .236$ for actions. However, post hoc Tukey HSD tests showed that the pattern of age effects was different for different types of detail. In terms of the accuracy of narrative details about both persons and objects, 2-year-olds were proportionately less accurate about person details than children in the three oldest groups, who in turn did not differ from each other. In addition, 3- and 4-year-olds were proportionately less accurate about person details than 12- and 13-year-olds. For narrative details about attributes, children in the two youngest groups did not differ from each other, although they were proportionately less accurate than children in all three older groups. Furthermore, 5- and 6-year-olds were proportionately less accurate than 12- and 13-year-olds, and children in the two oldest groups did not differ. In terms of details about location, 2-year-olds were significantly less accurate about location than children in the three oldest groups, which did not differ from each other, but 3- and 4-year-olds differed from no other group. Finally, in terms of children's accuracy about actions, children in the youngest two groups differed from those in the oldest two groups, children in the middle age group differed only from those in the oldest group, and children in the two oldest groups did not differ from each other. In terms of the interaction between

Table 6
Accuracy of types of narrative details: Means (and standard deviations) for the percentage accuracy of different content during initial, 1-, and 2-year follow-up interviews

Age (years)	Interview	Type of detail				
		Persons	Objects	Attributes	Locations	Activities
2	Initial	60.6 (34.4)	69.4 (36.4)	52.1 (46.5)	81.8 (36.5)	81.3 (22.0)
	1-Year	65.0 (21.7)	71.1 (31.8)	62.9 (37.0)	75.3 (39.3)	71.0 (28.9)
	2-Year	62.6 (24.6)	60.5 (22.4)	53.4 (31.9)	67.6 (38.7)	70.0 (27.6)
3 & 4	Initial	82.0 (21.2)	84.3 (21.4)	66.8 (30.7)	86.8 (25.2)	85.1 (18.0)
	1-Year	72.4 (22.6)	74.0 (23.9)	59.7 (27.2)	82.8 (25.9)	74.1 (22.4)
	2-Year	71.2 (23.4)	72.2 (19.2)	59.8 (18.2)	88.2 (21.2)	74.2 (19.1)
5 & 6	Initial	86.6 (18.6)	88.0 (15.6)	77.0 (21.7)	93.9 (13.3)	87.6 (16.6)
	1-Year	80.9 (22.0)	80.8 (22.2)	66.8 (18.1)	87.9 (20.9)	80.9 (18.3)
	2-Year	76.4 (22.5)	78.8 (21.0)	64.8 (16.2)	91.2 (20.1)	78.8 (17.7)
8 & 9	Initial	90.6 (16.0)	95.8 (7.3)	88.7 (20.0)	96.7 (11.4)	95.2 (9.0)
	1-Year	80.5 (17.7)	86.2 (17.2)	73.2 (17.5)	94.8 (13.1)	83.5 (18.5)
	2-Year	73.4 (22.4)	77.6 (21.1)	73.3 (22.4)	91.5 (12.4)	83.2 (17.8)
12 & 13	Initial	92.7 (15.4)	95.1 (13.4)	93.7 (10.3)	98.3 (5.6)	97.2 (6.0)
	1-Year	84.0 (19.3)	88.3 (15.8)	80.0 (15.9)	96.7 (6.8)	93.7 (8.2)
	2-Year	84.2 (15.5)	80.6 (21.3)	80.4 (14.8)	96.3 (6.8)	94.3 (6.4)
All	Initial	83.6 (22.5)	86.6 (21.0)	74.4 (29.6)	91.2 (21.0)	88.2 (16.8)
	1-Year	76.8 (22.0)	79.1 (23.3)	66.4 (24.3)	86.8 (24.0)	79.5 (21.1)
	2-Year	74.0 (22.6)	74.9 (21.3)	65.5 (21.2)	88.3 (22.5)	79.3 (19.0)

detail type and gender, the main effect of gender was nonsignificant except for details about attributes, $F(1, 135) = 5.22$, $p = .024$, $\eta^2 = .037$. Girls ($M = 70.0\%$) were more accurate about attributes than boys ($M = 66.9\%$). There were no interactions between gender and other factors for any type of detail.

To summarize, older children had more complete, informative, and accurate recall than younger children, and gender was mostly negligible, although there were occasional small effects favoring girls. Of more interest is how children's reports changed over time. The completeness of their reports did not change, whereas the amount of narrative detail increased over time for every type of detail and especially for actions and attributes. Only accuracy deteriorated over time, with children making proportionately more commission errors, and this was true for both completeness components and narrative details. In addition, central completeness components were recalled better than peripheral ones.

Discussion

When children are recalling a highly salient and stressful event, the completeness of their recall in terms of components of the overall event, the amount of narrative detail they provide, and the accuracy of their reports seem to follow different paths over time. Although common wisdom is that the quality of children's reports deteriorates over time, the current study suggests that it all depends on what aspect of those reports is being considered. Although accuracy (i.e., the proportion of information identified as commission errors) indeed is worse as the delay between event occurrence and recollection is greater, this is not true for children's recall of the structural components (i.e., completeness) of the event, which remains unchanged over time. Nor is it true for the amount of elaborative detail that children provide, which increases. This investigation was motivated by the need to shed light on the divergent conclusions that different investigations have drawn about children's long-term memory for highly salient events, and it appears that part of that explanation may be due to different researchers coding different aspects of children's memory reports.

All of these measures code different, yet important, aspects of children's recall. In terms of how well children provide information relevant to the overall structure of the event (i.e., information about where, who, what happened, etc.), they recalled on average three-quarters of those components identified as central to their injury experience and two-thirds of those components identified as peripheral, for a total recall completeness of approximately 73%. The current study extended (using a larger sample size) previous studies that reported little change in recall completeness for injury events (Peterson, 1999; Peterson & Bell, 1996; Peterson & Whalen, 2001). The fact that central components were recalled better than peripheral ones also replicates other research investigating information centrality (Goodman et al., 1991; Peters, 1997; Peterson & Bell, 1996; Vandermaas, Hess, & Baker-Ward, 1993). However, as one would expect when such a large age range is investigated, there were substantial differences in completeness depending on children's age. The 2-year-olds recalled only approximately half of the relevant components of their injury, whereas the older children recalled more. But importantly, there was little decrease in completeness over the 2 years investigated here. These events were so salient that they continued to be well recalled even 2 years later.

In contrast to children's reports remaining equivalently complete across time, they contained more narrative detail in subsequent interviews. By their 2-year follow-up interview, children provided on average 1 additional narrative detail about each of the categories of person, location, and object and provided 3 additional details about both attributes and actions. In other words, they provided a more detailed account for every type of information that had been scored, and this increase was particularly sharp between the 1- and 2-year follow-up interviews. Thus, as children's linguistic competence increased, it was reflected by greater narrative elaboration. It should be noted that an increase across age in narrative elaboration has been found for many years, but this is when children are studied cross-sectionally and, thus, children of different ages are providing initial reports of events. For example, in an early study, Peterson and McCabe (1983) found systematic increases in narrative structural sophistication and elaboration across the preschool and elementary years, and many reports since then have shown increasing competence across age in a host of narrative skills (see review in Nelson & Fivush, 2004), including the amount of elaborative detail they include. However, the current study suggests that such an increase in detail seems to be true not only for narratives that are first constructed at a particular age by different children but also for the *same* children's reports about the *same* events across time. This replicates (with a different scoring system and a different event) what Fivush et al. (2004) and Sales et al. (2005) reported about children's recall of a destructive hurricane. Parallel to the children recalling a hurricane, the children in the current investigation demonstrated reminiscence by providing new information in later interviews that had not been presented in earlier interviews. Such reminiscence effects have been frequently found over repeated interviews with both children and adults (La Rooy et al., 2005). Children also demonstrated hypermnesia given that the total amount of detail increased over time. Thus, as children's narrative competence increased, they provided more elaborative detail about events that they had recounted on previous occasions.

How can children maintain equivalent completeness scores over time and yet provide more information? The answer is related to how these two variables are scored. Recall the difference between "We were at my Nan's" and "We were at my Nan's by the green shed that's next to her house." These two descriptions (provided by the same child about where she was injured in her initial and 2-year follow-up interviews) are equivalent in terms of recall completeness given that they both provide information relevant to the prototype component "Where did the injury occur?" But the second description provides more narrative detail. It is possible that the increased narrative detail is due to parents talking about the events after each interview and, thus, children incorporating this additional information into their reports the next time they were interviewed. That is, children acquire postevent correct information through family discussions, and this could inflate recall performance. However, if this is the primary explanation for the increase in information, it is odd that the largest increase occurred not between the initial and second interviews (when the injury was most talked about according to parental report) but rather between the second and third interviews (when parents reported that very little if any discussion of these events from long ago took place).

Another possibility is that being asked direct questions in earlier interviews increases later recall for the aspects that are asked about. Such reinstatement effects are certainly well-known phenomena that can assist long-term recall. In addition, having a delayed earlier interview can lead to better long-term recall, probably due to memory reactivation being particularly effective after some forgetting has occurred (Pipe, Sutherland, Webster, Jones, & La Rooy, 2004). In terms of children's recall of injuries requiring emergency room treatment, the effect of multiple interviews on completeness of injury prototype components and their proportion accuracy were explored in two previous studies of 3- to 9-year-olds. The first compared (a) children who had three interviews, immediately after injury and again 6 months and 1 year later, (b) children who had two interviews, immediately after injury and 1 year later, and (c) children who had only one interview, 1 year after injury (Tizzard-Drover & Peterson, 2004). Having an immediate interview (regardless of whether there was an extra one at 6 months) resulted in 3- and 4-year-olds having more complete and more accurate recall in their 1-year interview, although there was no effect for older children. That is, having a detailed interview was beneficial for younger children's recall, probably because the early interview provided a comprehensive reinstatement of the event. At 2 years postinjury, there was no difference between groups in their recall of their injury regardless of the number of prior interviews children had experienced or whether their first interview had been 1 week or 1 year after injury (Peterson, Pardy, Tizzard-Drover, & Warren, 2005). However, neither of these studies assessed the number of narrative details that children provided, only the number of completeness components and the proportion of them that were correct. Thus, it is possible that the increase in the number of narrative details that children recall is related to having had earlier interviews to remind them.

A third possibility is that when children were interviewed 2 years later, they were more verbally competent and, consequently, were able to be more descriptive when talking about the same prototype components that they had told us about 2 years earlier. That is, they provided the same fundamental tale of what had happened to them but were now able to embellish it with more detail. In addition, they had learned more about cultural expectations for narratives and for how informative they were expected to be. In fact, considerable development in narrative skills is taking place over these ages (Nelson & Fivush, 2004), and such development is reflected in longer and more informative narratives. Yet another possibility is that children have an increased understanding of the social expectations surrounding an interview. That is, they have greater sensitivity to the perceived demands of an interview and know that they are expected to provide more detailed information.

The story is different, however, when accuracy is considered. In this respect, children's reports clearly get worse. That is, they make more commission errors over time. This is true for the accuracy of components used in the scoring of completeness and for the accuracy of narrative details. However, one should not lose sight of the substantial overall accuracy of children's reports even 2 years later. For the proportion accuracy of completeness components, accuracy slipped from approximately 92% to 85%, and for the proportion accuracy of narrative detail, accuracy decreased from approximately 92% to 86%. A decrease in accuracy and an increase in commission errors across time have been found in many studies (see the meta-analytic review by Deffenbacher, Bornstein, Penrod, & McGorty, 2004).

Interestingly, the decrease in accuracy across time was equivalent for central and peripheral components of report completeness, and the decrease in accuracy was also equivalent for all five types of detail content. However, that does not mean that all types of content were recalled with equivalent accuracy. For younger children only, peripheral components were recalled less accurately than central components, whereas older children recalled both types of components with equivalent accuracy. And for all children, attributes were recalled with considerably less accuracy (~69% of such details were correct) than were other types of details, especially those pertaining to locations (~89% were correct).

Gender had relatively little effect on children's recall. However, girls were more accurate than boys in terms of completeness components, although this difference was very modest (88.9% vs. 87.0%). Thus, although it was statistically significant, it had little practical importance. In addition, there were some gender differences associated with specific types of narrative detail. Girls were more accurate than boys when providing narrative details about attributes (70.0% vs. 66.9%). In addition, girls provided more narrative details about people as their age increased, whereas boys of all ages continued to provide the same number of person details. Previous research has found that girls are likely to refer to more people in their narratives than are boys (Buckner & Fivush, 1998, 2000) and that mothers are likely to place events within a more interpersonal context when talking with their daughters than when talking with their sons (Fivush, Berlin, Sales, Mennuti-Washburn, & Cassidy, 2003). Thus, the modest gender differences that were found here are consistent with findings of other research that was quite different in methods and narrative topics.

To summarize, children's recall completeness, provision of narrative detail, and accuracy seem to follow different paths over time in children's reports about a highly salient event. Narrative details increase, whereas accuracy decreases, and children continue to provide the same overall structure when describing the event (i.e., are equivalently complete). A focus by different investigators on different aspects of children's memory reports about highly salient events, along with the use of different coding schemes, may help to explain why there are such discrepancies in conclusions between different investigators about what happens to children's memory reports over time. In terms of forensic implications, the general belief that children's reports get worse over time is too simplistic. It is only the accuracy of their reports that deteriorates, and the current study suggests that the deterioration is quite modest even for young children. In contrast, as children get older, they increasingly come to understand that interviewers both want and expect them to provide a lot of information about events and, in particular, to include in their reports information about who, what, where, when, how, and why. They also have better linguistic and cognitive skills that enable them to be more elaborative in their reports. One consequence of this developing understanding of narrative structure and cultural expectations is that children's free recalls (the traditional beginning of a forensic interview) get longer and more informative over time. Even when interviewers turn to prompting questions, children understand that more is better when it comes to providing information. Also comforting is the fact that children's understanding of the overall event – its structure and components – does not deteriorate either. It is only details here and there that deteriorate. And this deterioration is across the board in terms of the type of information (persons, objects, attributes, and activities). The one exception is that details about location seem to be forgotten less than other types of details.

There are limitations in the current study, and foremost is the fact that we were unable to interview the children until several days had passed. It is unfortunate that children could not have been interviewed before there had been rehearsal of the events with relatives and friends, but ethical considerations when studying naturally occurring events leave little choice. After all, the events were big news that, of course, would have been widely discussed, and ethically one cannot interview families until they have had an opportunity to calmly consider whether they wish to participate in research.

Nor did we have any control over how much these events were talked about between interviews, although parents reported that after children were healed these events became "old news" and were seldom if ever mentioned. In fact, parents were typically quite surprised at how well their children recalled the events 2 years later and said that they had not talked about the events since the last time we had visited if even then. In addition, not all of the children's details were able to be confirmed, and so it is possible that unconfirmed details were not distributed in the same way that confirmed details were distributed in terms of accuracy. In addition, reminisced information is vulnerable to errors (Gilbert & Fisher, 2006; La Rooy et al., 2005). Nevertheless, the fact that the sample size was so large suggests that the findings are likely to be robust.

In conclusion, children can become better witnesses over time by providing more elaborative detail about events, and this is an important contributor to believability (Bala et al., 2001). They also become worse witnesses over time by making more commission errors, which is an important contributor to disbelief of their stories. Knowing how children's reports are expected to change can potentially help professionals to assess the credibility of child witnesses.

Acknowledgments

This research was primarily supported by Grant 513-02 from the Natural Sciences and Engineering Research Council of Canada. Additional funding came from the Memorial University Undergraduate Career Experience Program and the Student Work and Service Program. I thank all of the recruiters, interviewers, transcribers, and data analyzers who participated and am especially grateful to Penny Voutier for her data management and Malcolm Grant for statistical consultation. I also thank the Janeway Hospital emergency room staff and, most of all, the parents and children who allowed us into their homes and cooperated so helpfully.

Appendix A

Prototype of interview about the injury with examples of items and classification as central or peripheral

Item	Example	Category
Time of day	"*Right after lunch*"	P
Place	"In my *backyard*"	P
Who was with you	"*Mom* and my *brother* Joe"	P
Who else was around	"My *friend Anna* was playing there too"	P
Actions prior to injury	"I was *running*"	P
The injury	"I got a big *cut* on my leg"	C
How it occurred	"I was *tripped*"	C
Who did it	"By my *brother*"	C
What objects involved	"I hit a *piece of the porch* that was sticking up"	C
Cry	"I had to just *scream*"	C
Blood	"It was *bleeding* all down my leg"	C
Who first responded	"*Mommy* heard me cry"	C
Where you went before hospital	"She took me into the *kitchen*"	P
Actions to treat injury	"She *wiped* my knee"	C
Objects of home treatment	"And put a *cloth* on my knee to soak up blood"	C
Anyone else look/help	"My *brother* was watching"	P
Went to hospital	"Then I went to the *hospital*"	C
Who took you to hospital	"*Mom* drove me there"	P
Who else went along	"My *brother* had to come too"	P
Time of hospital trip	"We got to the hospital *half an hour later*"	P

Note. C, central; P, peripheral.

References

Bahrick, L. E., Parker, J. F., Fivush, R., & Levitt, M. (1998). The effects of stress on young children's memory for a natural disaster. *Journal of Experimental Psychology: Applied, 4*, 308–331.

Baker-Ward, L., Gordon, B. N., Ornstein, P. A., Larus, D. M., & Clubb, P. A. (1993). Young children's long-term retention of a pediatric examination. *Child Development, 64*, 1519–1533.

Baker-Ward, L., Ornstein, P. A., Gordon, B. N., Follmer, A., & Clubb, P. A. (1995). How shall a thing be coded? In M. S. Zaragoza, J. R. Graham, G. C. N. Hall, R. Hirschman, & Y. S. Ben-Porath (Eds.), *Memory and testimony in the child witness* (pp. 61–85). Thousand Oaks, CA: Sage.

Bala, N., Lee, J., & McNamara, E. (2001). Children as witnesses: Understanding their capacities, needs, and experiences. *Journal of Social Distress and the Homeless, 10*, 41–68.

Bauer, P. J. (2006). Event memory. In D. Kuhn & R. S. Siegler (Eds.), *Handbook of child psychology: Cognition, perception, and language, Vol. 2* (6th ed., pp. 373–425). John Wiley: New York.

Bauer, P. J. (2007). *Remembering the times of our lives: Memory in infancy and beyond*. Mahwah, NJ: Lawrence Erlbaum.

Bluck, S., Levine, L. J., & Laulhere, T. M. (1999). Autobiographical remembering and hypermnesia: A comparison of older and younger adults. *Psychology and Aging, 14*, 671–682.

Boland, A. M., Haden, C. A., & Ornstein, P. A. (2003). Boosting children's memory by training mothers in the use of an elaborative conversational style as an event unfolds. *Journal of Cognition and Development, 4*, 39–65.

Bornstein, B. H., Liebel, L. M., & Scarberry, N. C. (1998). Repeated testing in eyewitness memory: A means to improve recall of a negative emotional event. *Applied Cognitive Psychology, 12*, 119–131.

Buckner, J. P., & Fivush, R. (1998). Gender and self in children's autobiographical narratives. *Applied Cognitive Psychology, 12*, 407–429.

Buckner, J. P., & Fivush, R. (2000). Gendered themes in family reminiscing. *Memory, 8*, 401–412.

Burgwyn-Bailes, E., Baker-Ward, L., Gordon, B. N., & Ornstein, P. A. (2001). Children's memory for emergency medical treatment after one year: The impact of individual difference variables on recall and suggestibility. *Applied Cognitive Psychology, 15*, 25–48.

Christianson, S.-A., & Loftus, E. F. (1991). Remembering emotional events: The fate of detailed information. *Cognition & Emotion, 5*, 81–108.

Deffenbacher, K. A., Bornstein, B. H., Penrod, S. D., & McGorty, E. K. (2004). A meta-analytic review of the effects of high stress on eyewitness memory. *Law and Human Behavior, 28*, 687–706.

Fivush, R. (1991). The social construction of personal narratives. *Merrill-Palmer Quarterly, 37*, 59–82.

Fivush, R., Berlin, L., Sales, J. M., Mennuti-Washburn, J., & Cassidy, J. (2003). Functions of parent–child reminiscing about emotionally negative events. *Memory, 11*, 179–192.

Fivush, R., Sales, J. M., Goldberg, A., Bahrick, L., & Parker, J. (2004). Weathering the storm: Children's long-term recall of Hurricane Andrew. *Memory, 12*, 104–118.

Gilbert, J. A. E., & Fisher, R. P. (2006). The effects of varied retrieval cues on reminiscence in eyewitness memory. *Applied Cognitive Psychology, 20*, 723–739.

Goodman, G. S., Hirschman, J. E., Hepps, D., & Rudy, L. (1991). Children's memory for stressful events. *Merrill-Palmer Quarterly, 37*, 109–158.

Goodman, G. S., Quas, J. A., Batterman-Faunce, J. M., Riddlesberger, M. M., & Kuhn, J. (1994). Predictors of accurate and inaccurate memories of traumatic events experienced in childhood. *Consciousness and Cognition, 3*, 269–294.

Goodman, G. S., Quas, J. A., Batterman-Faunce, J. M., Riddlesberger, M. M., & Kuhn, J. (1997). Children's reactions to and memory for a stressful event: Influences of age, anatomical dolls, knowledge, and parental attachment. *Applied Developmental Science, 1*, 54–75.

Haden, C. A., Ornstein, P. A., Eckerman, C. O., & Didow, S. M. (2001). Mother–child conversational interactions as events unfold: Linkages to subsequent remembering. *Child Development, 72*, 1016–1031.

Heuer, F., & Reisberg, D. (1992). Emotion, arousal, and memory for detail. In S. A. Christianson (Ed.), *Handbook of emotion and memory* (pp. 151–180). Hillsdale, NJ: Lawrence Erlbaum.

La Rooy, D., Pipe, M.-E., & Murray, J. E. (2005). Reminiscence and hypermnesia in children's eyewitness memory. *Journal of Experimental Child Psychology, 29*, 235–254.

Merritt, K. A., Ornstein, P. A., & Spicker, B. (1994). Children's memory for a salient medical procedure: Implications for testimony. *Pediatrics, 94*, 17–23.

Nelson, K., & Fivush, R. (2004). The emergence of autobiographical memory: A social cultural developmental theory. *Psychological Review, 111*, 486–511.

Ornstein, P. A., Haden, C. A., & Hedrick, A. M. (2004). Learning to remember: Social-communicative exchanges and the development of children's memory skills. *Developmental Review, 24*, 374–395.

Peters, D. P. (1997). Stress, arousal, and children's eyewitness memory. In N. L. Stein, P. A. Ornstein, B. Tversky, & C. Brainerd (Eds.), *Memory for everyday and emotional events* (pp. 351–370). Mahwah, NJ: Lawrence Erlbaum.

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

Peterson, C. (2002). Children's long-term memory for autobiographical events. *Developmental Review, 22*, 370–402.

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

Peterson, C., & McCabe, A. (1983). *Developmental psycholinguistics: Three ways of looking at a child's narrative*. New York: Plenum.

Peterson, C., Pardy, L., Tizzard-Drover, T., & Warren, K. (2005). When initial interviews are delayed a year: Effect on children's 2-year recall. *Law and Human Behavior, 29*, 527–541.

Peterson, C., & Roberts, C. (2003). Like mother, like daughter: Similarities in narrative style. *Developmental Psychology, 39*, 551–562.

Peterson, C., & Whalen, N. (2001). Five years later: Children's memory for medical emergencies. *Applied Cognitive Psychology, 15*, 7–24.

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

Quas, J. A., Bauer, A., & Boyce, W. T. (2004). Physiological reactivity, social support, and memory in early childhood. *Child Development, 75*, 797–814.

Quas, J. A., Goodman, G. S., Bidrose, S., Pipe, M.-E., Craw, S., & Ablin, D. S. (1999). Emotion and memory: Children's long-term remembering, forgetting, and suggestibility. *Journal of Experimental Child Psychology, 72*, 235–270.

Sales, J. M., Fivush, R., Parker, J., & Bahrick, L. (2005). Stressing memory: Long-term relations among children's stress, recall, and psychological outcome following Hurricane Andrew. *Journal of Cognition and Development, 6*, 529–545.

Salmon, K., Price, M., & Pereira, J. K. (2002). Factors associated with young children's long-term recall of an invasive medical procedure: A preliminary investigation. *Developmental and Behavioral Pediatrics, 23*, 347–352.

Shrimpton, S., Oates, K., & Hayes, S. (1998). Children's memory of events: Effects of stress, age, time delay, and location of interview. *Applied Cognitive Psychology, 12*, 133–143.

Tizzard-Drover, T., & Peterson, C. (2004). The influence of an early interview on long-term recall: A comparative analysis. *Applied Cognitive Psychology, 18*, 1–17.

Vandermaas, M. O., Hess, T. M., & Baker-Ward, L. (1993). Does anxiety affect children's reports of memory for a stressful event? *Journal of Applied Psychology, 7*, 109–128.