APPLIED COGNITIVE PSYCHOLOGY
Appl. Cognit. Psychol. **18**: 613–631 (2004)
Published online in Wiley InterScience
(www.interscience.wiley.com) DOI: 10.1002/acp.1013

Timing Moderates the Effects of Repeated Suggestive Interviewing on Children's Eyewitness Memory

LAURA MELNYK[1]* and MAGGIE BRUCK[2]

[1]*King's College, University of Western Ontario, Canada*
[2]*Johns Hopkins University, USA*

SUMMARY

The relative role of the timing and repetition of misinformation on the accuracy of children's recall was examined in two experiments. Kindergarten children participated in a magic show and about 40 days later had a memory test. Between the magic show and the memory test, the children were suggestively interviewed either one time in a relatively 'early' interview (temporally closer to the magic show than the memory test) or a relatively 'late' interview (closer to the memory test than the magic show), or in both suggestive interviews. The timing of the suggestive interviewing was manipulated so that the interview was temporally distant from the event or memory test or temporally close to the event or memory test. Repeated interviewing heightened misinformation effects only when the children received the two interview sessions temporally close to the event and memory test. Copyright © 2004 John Wiley & Sons, Ltd.

A basic principle of memory is that repetition promotes recall (Ebbinghaus, 1885/1913). The research to support this principle concerns the beneficial effects of repeating accurate information on participants' recall. There is however, little direct evidence to examine the corollary of this principle, namely the repetition of erroneous information (misinformation) promotes memory distortion. The present study examines this hypothesis with respect to one particular type of memory distortion, false reports that are the product of suggestive interviews in which misinformation has been presented. It examines the cumulative effect of repeated misinformation on memory distortion in young children.

Repeated neutral interviews (where there is no misinformation) may confer a number of positive effects on recall. A review of the developmental literature reveals that it can consolidate memory for an event and thus protect against forgetting (Baker-Ward, Hess, & Flannagan, 1990; Dent & Stephenson, 1979); it prevents normally occurring errors of commission (Dent & Stephenson, 1979; Hudson, 1990); and it may promote *reminiscence*, the reporting of previously unmentioned details in later recall (e.g. Howe, Kelland, Bryant-Brown, & Clark, 1992; Howe, O'Sullivan, & Marche, 1992, but see Cassel & Bjorklund, 1995). A final potential benefit of repeated interviewing is *hypermnesia*, an

*Correspondence to: Dr L. Melnyk, Department of Psychology, King's College, 266 Epworth Avenue, London, ON, N6A 2M3, Canada. E-mail: laura@uwo.ca

Contract/grant sponsors: McGill University; National Sciences and Engineering Research Council.

increase in the amount of new information recalled over increasing retention intervals that exceeds the amount of information forgotten (Memon, Wark, Bull, & Koehnken, 1997).

It is important to point out, however, that some studies did not find benefits of repeated interviewing on children's recall. For example, Poole and White (1991) found that although adults who were interviewed both immediately after witnessing a live event and 1 week later reported more information than adults who were only tested at the 1-week session, repeated interviewing did not influence the amount of information reported by 4- to 8-year-old children. In another study, the reports provided by 3-, 5-, and 7-year-old children 3 weeks after a doctor's examination were similar whether the children were interviewed both immediately after the exam and 1 week later or if the children were interviewed only after the 3-week delay (Baker-Ward, Gordon, Ornstein, Larus, & Clubb, 1993). Furthermore, in some cases repeated interviewing may decrease the accuracy of later recall by increasing the number of errors reported (Fivush & Hamond, 1989).

The inconsistent results in this literature may be partly accounted for by the specific timing of the neutral interview. A window of time may limit the positive effects of a neutral interview on later recall. Benefits may only occur when the neutral interview is within a period of about 1 week, and not several weeks, after the event (e.g. Powell & Thomson, 1997).

We now turn to studies on the effects of repeated interviews that contain misinformation. In a misinformation paradigm, participants experience or witness an event, are given some misinformation about the details of the event, and are later asked to recall the original event. The number of misleading details reported serve as the index of the participants' susceptibility to misinformation about the event (e.g. Ceci, Ross, & Toglia, 1987; Cohen & Harnick, 1980; Goodman & Reed, 1986; Loftus, 1975). Although a number of studies demonstrate that repeated suggestive interviews taint the accuracy of children's memories (e.g. Bruck, Ceci, Francoeur, & Barr, 1995; Bruck, Ceci, & Hembrooke, 2002; Leichtman & Ceci, 1995), these studies have not systematically examined (a) how the *timing* of the misinformation influences false reporting or (b) how the *repetition* of misleading information influences false reporting. For example, if misinformation effects are not obtained after one suggestive interview, but are obtained after two interviews, it is not clear if it is the number of interviews that culminated in false reports or whether memory was significantly weakened by the second interview. The independence of these two factors must be kept in mind when evaluating the existing literature.

Warren and Lane (1995) found that repetition of misinformation did not negatively affect the accuracy of 9-year-olds' reports. Interviewing shortly after presentation of a video, regardless of whether the interview was neutral or misleading, improved children's recall at a 1-week memory test. Warren and Lane concluded that an early interview, even if it contains some misleading information, helps to consolidate memory for the event and provides protection against forgetting. In addition, it was found that while early neutral interviewing was more effective than later neutral interviewing for increasing recall accuracy, multiple suggestive interviews may not further influence suggestibility and therefore the timing of misinformation, not the repetition of misinformation, is a more influential factor on children's recall accuracy. Other studies suggest that the timing of misinformation influences children's suggestibility in much the same way as adults' suggestibility, with misinformation posing the greatest risk to accurate recall when it is temporally further from the event and closer to the memory test (Gobbo, 2000; Tucker, Mertin, & Luszcz, 1990). In Gobbo's study (2000, experiment 2), however, because all children who received the late misinformation also received two doses of misinformation, the results could reflect repetition rather than timing effects.

Roberts, Lamb, and Sternberg (1999) did not replicate the timing of misinformation effects, but their discrepant results may in fact shed light on the boundaries in which such effects are found. The timing of a suggestive interview (1 day or 4 weeks before a memory test occurring 5 weeks after a costume session) influenced the accuracy of 4-year-olds' reports in the memory test; children who had the 4-week interview produced more accurate details in free recall than children who had the 1-day interview or no interview. However, the timing of the interview did not influence the strength of the misinformation effects. Roberts et al.'s failure to find an effect for timing of misinformation may reflect the fact that there was a 1-week delay between the presentation of the misinformation at 1 month and the memory testing. In contrast, Gobbo (2000) and Warren and Lane (1995) presented the misinformation for their 'late' interview groups immediately before the memory test. Thus the results obtained by Gobbo (2000) and Warren and Lane (1995) may be due to the relative 'strength' of the misinformation whereas the week-long delay employed by Roberts et al. may have reduced the relative strength of the misinformation. In addition, as Roberts et al. acknowledged, the low rates of misinformation acceptance obtained in their study may have constrained their results.

Marche (1999) attempted to separate the components of timing and repetition by examining the relationship between the strength of the memory trace for the target event and repeated presentations of misinformation. Consistent with earlier studies (Marche & Howe, 1995; Pezdek & Roe, 1995), children with weaker memory traces of the initial event (that is, those who saw the slide show once) reported more misinformation than the children with stronger memory traces (children who saw the slide show multiple times). In addition, repetition of the misinformation had differential effects which were linked to the strength of the original memory: Children exposed to the event multiple times were susceptible only to repeated presentations of misinformation while the children who saw the event only once were equally susceptible to the misinformation, regardless of the number of presentations. Thus, for children who experienced the event once, the effects of repeated presentations of the misinformation was no more detrimental than a single one. As for the issue of timing, contrary to results described previously, there was no evidence that the misinformation given closer in time to the memory test increased the misinformation effects.

Given this limited literature, there is little consistent data on the effects of repeating (mis)information about experienced events on the accuracy of children's reports. There are hints that the effects of repetition may be mediated by the timing of the interviews relative to the target event and to the memory test. Thus it may not be a matter of one or more suggestive interviews, but when each of these interviews occurs. For example, if the placement of the first interview is close to the event (as typically occurs in the existing research), additional presentations of the misinformation do not increase children's overall suggestibility (e.g. Gobbo, 2000; Warren & Lane, 1995). But if the first interview was given after a longer delay, would a second interview increase misinformation effects? In forensic settings, children are questioned repeatedly about situations that happened weeks, months, and sometimes years earlier (e.g. Ceci & Bruck, 1995).

A second question addresses *how* the timing and repetition of misinformation independently influence the accuracy of children's recall. The existing literature does not include a simple test which compares children's recall facilitation and misinformation effects when leading and misleading suggestions are presented (1) once, temporally close to the event, (2) once, close to the memory test, or (3) twice, at the same intervals as groups 1 and 2. In the present paper, Experiment 1 examined the relative role of the repetition and

timing of suggestive interviews. The suggestive interviews contained both accurate information (correctly leading suggestions) and misinformation (misleading suggestions) about an event that the children had experienced. Children were interviewed either once in a relatively 'early' or 'late' suggestive interview, or twice, in both the early and late suggestive interviews. In order to selectively examine the effects of *repetition* on children's recall accuracy, the timing of the 'early' interview was conducted outside of the 1-week time frame associated with consolidation (e.g. Powell & Thomson, 1997; Warren & Lane, 1995). Because Gobbo (2000) and Warren and Lane (1995) conducted their 'early' interview immediately after the event, their design may have masked the effect of repetition because the early interview consolidated memory for the event. In addition, since a 'late' interview immediately before a memory test may increase suggestibility because of the relatively strong memory trace for the misinformation, the 'late' interview in Experiment 1 was conducted about 2 weeks before the memory test to minimize the relative strength of the misinformation and to allow a more specific focus on the repetition of events.

If the *repetition* of (mis)information alone is sufficient to heighten facilitation effects and misinformation effects on a later memory test, then the children who had *both* suggestive interviews would show relatively greater facilitation effects and misinformation effects than the children who had only the early or late suggestive interview, who would not differ from each other on the memory test. However, if the *timing* of the suggestive interviews is critical to heighten facilitation effects and misinformation effects, then repetition effects would not differ between the children who received one or two suggestive interviews.

A related issue with theoretical and applied importance involves the long-term consequences of suggestive interviewing. Although most studies of children's suggestibility involve delays of a few minutes to several weeks between the presentation of misinformation and the memory test (e.g. Ackil & Zaragoza, 1995, 1998; Bruck, Ceci, & Francouer, 2000; Bruck, Melnyk, & Ceci, 2000; Ceci et al., 1987; Marche, 1999; Marche & Howe, 1995; Roberts et al., 1999), in the legal arena, delays of several months often intervene between interviews with children (where there may be misleading suggestions) and the start of a trial (Flin, 1995; Goodman et al., 1992), which may be viewed as the final memory test. Therefore, an important question is the degree to which misinformation affects the long-term accuracy of children's reports.

The only two published studies examining the fate of suggested reports over several weeks or years indicate that reports of false events and misinformation effects decrease substantially over time. Ceci, Huffman, Smith, and Loftus (1994) repeatedly interviewed preschoolers about true and false autobiographical events. Two years later, Huffman, Crossman, and Ceci (1997) re-interviewed a subset of the children about these events. The children's accuracy in correctly reporting the true events remained high but the rate of false reporting declined. Huffman et al. reasoned that the children's relatively higher recall for the true events may have been due to greater rehearsal and elaboration for these events compared to the false events because the true events involved special family occasions.

In a study examining the fate of misinformation effects, Poole and Lindsay (2001) reported that the frequency of 3- to 8-year-old children's false reports decreased over a period of several weeks. Misinformation effects were weaker in a second memory test about 4.5 months later compared with an initial test at 3.5 months, with the children reporting fewer false details in response to the open-ended prompts and fewer false details in response to the yes/no questions. Poole and Lindsay reasoned that false reports should

decline more sharply than accurate reports because the false details were more recently presented, and forgetting follows a negatively accelerated curve, with most recent information forgotten at a faster rate than older information (see also a review by Brainerd & Poole, 1997). Despite the fact that false reports declined in both studies, they nonetheless persisted.

In the present study, we examined the degree to which suggestions (leading and misleading) that were presented once or twice affected the accuracy of children's recall approximately 6 months later. We used preschool children because of concerns about the accuracy of preschoolers' eyewitness recall and because more reliable misinformation effects have been found in preschoolers compared to older children and adults (see Ceci & Bruck, 1995).

EXPERIMENT 1

Method

Design

Pairs of preschool children interacted with a magician who performed two tricks. Some children participated in suggestive interviews approximately 14 days and 25 days after the magic show. A control group did not participate in the suggestive interviews. An initial memory test was conducted approximately 40 days after the magic show; all of the children provided free recall of the magic show and completed a probed recognition task. Approximately 21 weeks after the initial memory test, all of the children had a follow-up memory test, in which the initial memory test procedures were repeated.

Participants

One-hundred and thirteen children attending kindergarten programmes were recruited from six schools in the Montreal area. The children were between the ages of 60 and 73 months and were predominantly from middle-class socioeconomic backgrounds. All children had written parental permission to participate.

The children were randomly assigned to one of four timing conditions: the early suggestive interview condition ($n = 29$), the late suggestive interview condition ($n = 28$), the two suggestive interviews condition ($n = 28$), or the no suggestive interview control condition ($n = 28$). The groups were equated in terms of age (the mean age for each group was 68 months) and gender (between 41 and 54% of the children in each condition were girls).

Procedure

Session 1: magic show. A female experimenter visited the classroom and took pairs of children to another room in the school. To make the children comfortable, she engaged the children in conversation while playing with a sticker book. After a few minutes, the experimenter performed a scripted magic show that included the following events (the 20 pre-specified target details are in italics):

> The *magician put on a black hat* and *asked each child to put on a red magic helper cape*. She *told the children that 'abracadabra' was the magic word* that they would say for the tricks; the children and the magician practiced saying this word. The *magician took a stuffed toy rabbit out of her magic box* and put the rabbit on the table beside her magic stage. The magician performed two tricks. In the first trick, the *magician put a red ball in her pocket*. The *magician waved her wand* and the ball reappeared in a small blue

vase. The *magician covered her hand with a magic scarf* and each *child honked a dinosaur horn*, and the red ball reappeared in her pocket. In the second trick, each *child poured some water into a cup*. The *magician poured the water from the cup into a bowl*. The *magician stirred a 'magic' cup with a spoon* and tipped the cup over but nothing came out of the cup.[1] *The magician put the magic cup in the corner* and said, 'Stay, magic cup, stay!' After the trick, the *magician tripped over her shoelaces and laughed as she fell down*, saying how clumsy and silly she was. The *magician said that she hurt her hand*, and *she put two plastic bandages on her hand*. At clean-up time, the *magician put her tricks away in a box*, and *brushed her magic stage* with a little brush. *She showed the children a picture of her new puppy* and *she gave each child an award for being such a good helper.*

Early suggestive interview. Approximately 2 weeks after the magic show ($M = 14.5$ days, $SD = 1.6$), a female experimenter, who was not the magician, individually interviewed the children in the early suggestive interview and two suggestive interviews conditions. Each session was tape- and/or video-recorded. There were two parts to the interview: warm-up phase and suggestion phase.

Warm-up. The interviewer told the children, 'I am the story lady. I like to tell stories with children. I brought my tape recorder so I could share the stories with my teachers and friends.' The interviewer turned the tape recorder on and asked the children to tell her about their siblings and school activities.

Suggestion phase. The 20 pre-specified target details in the magician event were divided into five scenes (setting up, ball trick, water trick, the fall, and clean-up). As shown in the Appendix, each correctly leading (true) target detail was paired with a generated misleading (false) detail. For example, the leading target detail 'The magician laughed when she fell' was paired with the misleading detail 'The magician cried when she fell.'

Each child was given five leading suggestions and five misleading suggestions. For each child, one leading suggestion and one misleading suggestion were taken from each of the five scenes. A leading suggestion and its paired misleading suggestion were never presented together; thus the same child would never receive the leading suggestion 'The magician laughed when she fell' and the misleading suggestion 'The magician cried when she fell.' In order to ensure that all details were equally used and that different combinations of leading and misleading suggestions were used across participants, 24 different suggestive interview versions were selected from a large pool of possible choices, such that each leading and misleading suggestion was used an equal number of times across interview versions. (As described below, the remaining 10 details, not used as suggestions, later served as non-suggested 'control' details on the recognition test, such that each detail was used an equal number of times across interviewing versions as a suggested and control detail.) The children were randomly assigned to the interview versions such that the same interview versions were used across the timing conditions.

The interviewer started the suggestion phase by telling the child that she had heard about some of the things that happened with the magician and that she wanted to talk about these things for their story. For example, the interviewer said, 'I heard that you wore red magic helper boots. Let's talk about that.' The interviewer paused and then repeated the suggestion and asked the child the question that was paired with the suggestion (i.e. 'Were

[1]A powder called 'Lightning Gel' caused the water to congeal on contact.

the boots too big for you?'). If a child resisted a suggestion, the interviewer repeated the suggestion and said, 'Well, let's just talk about that anyway.' The remaining leading and misleading suggestions were randomly presented in the same manner. The interviewer made comments throughout the suggestive interview, such as 'You are such a good storyteller' and 'You are doing a great job helping me.' Once all the suggestions had been presented, the interviewer slowly repeated each suggestion, preceded by 'You told me that . . . ' or 'You remembered that . . . .'

Late suggestive interview. Approximately 11 days after the early suggestive interview ($M = 11$ days, $SD = 2.2$), the same experimenter returned to interview children in the late suggestive interview and two suggestive interviews conditions. The same procedures were used as described for the early suggestive interview. The same details were suggested and resistance to the misleading suggestions was dealt with in the same manner as in the early suggestive interview.

Initial memory test. Approximately 2 weeks ($M = 14.3$ days, $SD = 2.4$) after the late suggestive interview (that is, about 40 days after the magic show), the initial memory test was conducted by a new female experimenter. The children's recall of the magic show was assessed through free recall questions and a recognition test.

Free recall of the magic show. The experimenter told the child that she was not present for the magician's visit. She asked the child to tell her everything that had happened during the magician's visit. After the child gave a response, the experimenter asked if the child could tell her 'one more thing.' This continued until the child could not report any more events.

Scoring. The number of utterances produced in the child's free recall was counted. An utterance was defined as a statement, bound by pauses, that contained one verb. For example, the statement, 'The magician wore a black hat and I wore a red magic helper cape' contains two utterances: (1) The magician wore a black hat and (2) I wore a red magic helper cape. Each utterance was categorized in terms of accuracy and also whether it had been suggested during the suggestive interview.

Recognition test for target events. Each child was asked 20 yes/no questions. Five questions probed the leading suggestions (e.g. 'Did the magician show you a picture of her puppy?') and five questions probed the misleading suggestions (e.g. 'Did the magician give you a hug?'). In addition, the child was asked five questions about leading control details (true non-suggested details) and five questions about misleading control details (false non-suggested details) that had not been used as suggestions for that child. This combination of leading and misleading control details was determined by the composition of the specific interview version that was assigned for each child, creating 24 different interview versions. As was the case for the leading and misleading suggestions, a leading and misleading control detail was selected from each of the five scenes. Each control detail could not be from the same pair as any other detail in the recognition test. Thus for each of the five scenes there was one question about a leading suggestion, a misleading suggestion, a leading control detail, and a misleading control detail.

Note that because the children in the control condition did not have any suggestive interviews, all the questions for the children in the control condition were labelled 'control' questions. Thus the control group was asked questions about 10 leading control and 10 misleading control details.

The experimenter gave each child two practice questions to ensure that the child understood the task. If the child did not seem to understand, the experimenter repeated the instructions, and gave the child two more practice questions.

If the child spontaneously reported any of the details in the free recall phase, the child was automatically credited with a 'Yes' response for that detail, and that detail was not asked in the recognition test. For example, if the child's free recall included the misleading suggestion, 'The magician hurt her leg,' then the question, 'Did the magician hurt her leg?' was not asked, but the child was credited with a 'Yes' response for that detail.

Scoring. For all groups except the control group, accuracy was assessed by counting the number of assents to leading suggestions and leading control details and the number of denials to the misleading suggestions and misleading control details. For the control group, accuracy was assessed by counting the number of assents to the leading control details and the number of denials to the misleading control details.

Follow-up memory test. Approximately 27 weeks after the magic show ($M = 27.1$, $SD = 2.0$), and about 21 weeks after the initial memory test ($M = 21.3$ weeks, $SD = 2.1$), a new female experimenter conducted a follow-up memory test. The procedures in this memory test were identical to the procedures used in the initial memory test; that is, there was a free recall and a recognition test for the target details from the magic show.

The mean length of the first suggestive interview was 10.17 min ($SD = 1.20$) and 10.31 min ($SD = 1.10$) for the second suggestive interview. The mean length of the initial memory test was approximately 5 to 11 min. The interviewers in the initial and follow-up memory tests were unaware of the child's timing condition.

Results

Two sets of analyses were carried out for each of the major dependent variables. First the data for the children in the three different timing conditions are reported (early suggestive interview, late suggestive interview, two suggestive interviews). Second, the data for the children in the control condition are entered into the model. In this second analysis, because the children in the control group did not receive leading or misleading suggestions, only the leading and misleading control details are analysed.

Free recall

A $3 \times 2 \times 2 \times 2$ ANOVA with repeated measures was carried out on the number of utterances produced in free recall. The independent variables were timing condition (early suggestive interview vs. late suggestive interview vs. two suggestive interviews) as the between-group factor and accuracy of utterance (accurate vs. inaccurate), utterance status (suggested vs. non-suggested), and memory test (initial vs. follow-up) as the repeated factors. The analysis produced significant main effects of accuracy, $F(1, 82) = 199.38$, $p < 0.001$, and utterance status, $F(1, 82) = 251.84$, $p < 0.001$. Overall, the children produced more accurate ($M = 4.09$, $SD = 4.31$) than inaccurate ($M = 0.78$, $SD = 1.56$) utterances. There was a significant two-way interaction of Accuracy $\times$ Utterance Status, $F(1, 82) = 134.94$, $p < 0.001$. Planned comparisons of the interaction showed that the children's accurate recall contained more non-suggested details than suggested details (M suggested $= 1.71$, M non-suggested $= 7.01$), $p < 0.001$, but their inaccurate recall contained an equal amount of suggested and non-suggested details (M suggested $= 0.34$, M non-suggested $= 1.22$), $p < 0.001$. There was also a significant two-way interaction of

Table 1. Mean number of free recall utterances (with standard deviations) in Experiment 1

		Memory test					
		Initial			Follow-up		
		Timing condition			Timing condition		
Accuracy	Utterance status	Early suggestive interview	Late suggestive interview	Two suggestive interviews	Early suggestive interviews	Late suggestive interviews	Two suggestive interviews
Accurate	Suggested[a]	1.07 (0.70)	1.21 (0.96)	1.75 (1.38)	1.00 (0.93)	0.93 (0.86)	1.07 (0.90)
	Non-suggested	8.07 (4.19)	5.68 (3.31)	7.00 (4.36)	7.38 (3.44)	6.46 (4.73)	7.43 (5.76)
Inaccurate	Suggested[a]	0.14 (0.35)	0.25 (0.52)	0.82 (0.91)	0.17 (0.38)	0.29 (0.46)	0.39 (0.57)
	Non-suggested	0.83 (0.93)	1.14 (1.53)	0.79 (1.45)	1.45 (2.03)	1.89 (3.01)	1.25 (2.52)

[a]The maximum number of suggested details is 5.

Utterance Status × Memory Test, $F(1, 82) = 4.48$, $p < 0.05$. Planned comparisons of the interaction indicated that although the overall amount of information recalled was similar in both the initial and follow-up memory tests, the amount of *suggested* information decreased while the amount of non-suggested information increased between the initial and follow-up memory tests, $p < 0.001$. There were no main effects or interactions involving the timing condition variable. The data are presented in Table 1.

Total report

Accuracy of total report was assessed by summing the accurate responses produced in free recall with the accurate responses on the recognition test. The total number of accurate responses was analysed using a four-way repeated ANOVA with timing condition (early suggestive interview vs. late suggestive interview vs. two suggestive interviews) as the between-group factor and with accuracy (leading details vs. misleading details), detail status (suggested details vs. control details), and memory test (initial vs. follow-up) as the repeated factors. There were significant main effects of accuracy, $F(1, 82) = 89.51$, $p < 0.001$, detail status, $F(1, 82) = 12.06$, $p < 0.001$, and memory test, $F(1, 82) = 5.09$, $p < 0.05$. There were significant two-way interactions of Accuracy × Detail Status, $F(1, 82) = 162.30$, $p < 0.001$, and Accuracy × Memory Test, $F(1, 82) = 5.33$, $p < 0.05$. Finally, there was a significant three-way interaction of Accuracy × Detail Status × Memory Test, $F(1, 82) = 5.00$, $p < 0.05$. Importantly, there were no significant main effects or interactions involving the timing condition variable; thus, contrary to the hypothesis, neither the relative timing of the suggestive interviews nor the repetition of the suggestive interviews influenced the accuracy of the children's reports. The data are presented in Table 2.

Planned comparisons of the three-way interaction of Accuracy × Detail Status × Memory Test showed that, across timing conditions, there was no change in accuracy between the initial and follow-up memory tests for the leading suggested, leading control, or misleading suggested details. However, the children were more accurate in rejecting the misleading control details in the initial memory test ($M = 80\%$) than in the follow-up memory test ($M = 69\%$), $p < 0.001$. (Although the raw scores were used in the analyses, for clarity, the results are presented in terms of percentage correct—the raw scores divided by 5.)

The change on misleading control details necessarily affected a change in size of the misinformation effect from the initial to the follow-up memory test. Although there were

Table 2. Mean total report accuracy (with standard deviations) in Experiment 1

		Memory test					
		Initial			Follow-up		
		Timing condition			Timing condition		
Accuracy	Suggestion status	Early suggestive interview	Late suggestive interview	Two suggestive interviews	Early suggestive interview	Late suggestive interview	Two suggestive interviews
Leading	Suggestion	4.55 (0.63)	4.68 (0.55)	4.46 (0.88)	4.52 (0.74)	4.71 (0.60)	4.54 (0.74)
	Control	3.48 (1.09)	3.71 (1.12)	3.61 (1.17)	3.62 (1.08)	3.86 (1.04)	3.59 (0.97)
Misleading	Suggestion	2.38 (1.43)	1.82 (1.25)	2.18 (1.19)	2.35 (1.34)	1.93 (1.39)	1.96 (1.14)
	Control	4.07 (1.00)	3.50 (1.37)	4.14 (1.01)	3.45 (1.06)	3.32 (1.36)	3.54 (1.07)

Note: The maximum for each cell is 5.

significant misinformation effects at both testing periods (accuracy was significantly lower on misleading suggested details than on misleading control details), the misinformation effect was smaller at follow-up because of decreased accuracy on misleading control details but unchanged accuracy on the misleading suggested details. This change did not occur for the leading details (because there was no change in accuracy from the initial to follow-up memory tests). Here, children showed significant and similar facilitation effects in the initial and follow-up memory tests. In other words, suggested details were recalled more accurately than non-suggested details at the initial memory test and this advantage for suggested details remained 6 months later at the follow-up memory test.

Control details versus control group

The finding that leading suggestions increased accurate recall and that misleading suggestions increased recall errors is based on within-group comparisons of recall accuracy for suggested vs. control (non-suggested) details. Although this is a standard design to evaluate the effects of suggestion in misinformation studies, and although there is evidence that rehearsal only affects the rehearsed details (Bruck, Melnyk, & Ceci, 2000; Cassidy & DeLoache, 1995), it could still be argued that it is necessary to include a control group of children who received no suggestions between the magic show and the initial memory test to accurately evaluate the effects of leading and misleading suggestions. Specifically, it is possible that the children's recall of the control details was influenced by the suggestions during the suggestive interviews. Inclusion of a group of children who did not receive any suggestions provides an uncontaminated baseline measure of recall accuracy for the control details. Consequently, all of the following analyses compare the control (non-suggested) details of the suggestively interviewed timing groups (early suggestive interview, late suggestive interview, and two suggestive interviews groups) to the control group's performance on the control details. For the free recall analyses, therefore, all utterances that reflected suggestions given to the early suggestive interview, late suggestive interview, and two suggestive interviews groups were excluded. For the total report analyses, the control group's responses to the 10 leading and 10 misleading details were compared to the five leading and five misleading control details for the three suggestively interviewed timing groups. (Since all of the questions were control details for the control group, the total report analysis used proportional data when comparing the control group's performance with the performance of the other groups.)

Free recall. A 4 (condition) × 2 (accuracy) × 2 (interview) repeated ANOVA was performed to examine whether the suggestions influenced the children's recall for *non-suggested* utterances. As such, the utterances produced by the early suggestive interview, late suggestive interview, and two suggestive interviews groups that reflected the suggested details were excluded from the analysis. The analysis yielded a main effect of accuracy, $F(1, 109) = 248.07$, $p < 0.001$, and memory test, $F(1, 109) = 6.55$, $p < 0.01$, but the condition variable did not yield a main effect nor was it involved in any interactions. Thus, neither the number of accurate non-suggested utterances nor the accuracy of the utterances produced in free recall was influenced by whether or not the children had a suggestive interview.

Total report. Accuracy of total report was assessed by summing the accurate responses produced in free recall with the accurate responses on the recognition test. The scores used in this analysis were for non-suggested details only. For the early suggestive interview, late suggestive interview, and two suggestive interviews groups, only control (non-suggested) details were included in the analysis. The percentage of correct responses was analysed by dividing the leading control and misleading control scores by 5 (for the three timing groups) and by 10 (for the control group). A 4 (condition) × 2 (accuracy) × 2 (interview) repeated ANOVA was performed on the total report scores. There was a main effect of interview, $F(1, 109) = 10.53$, $p < 0.01$. Importantly, there were no significant effects involving the condition variable; thus recall of the non-suggested details by the children in the three timing groups was not influenced by the suggestive interviews in which other details were rehearsed. There was a significant interaction of Accuracy × Interview, $F(1, 109) = 20.06$, $p < 0.001$, which indicated that across conditions, the accuracy on the misleading control details decreased from the initial memory test to the follow-up memory test, while accuracy for the leading control details remained relatively constant between the initial and follow-up memory tests. Similar to the three timing groups, the control group was less accurate in rejecting the misleading control details in the follow-up memory test than in the initial memory test. In addition, like the three timing groups, there was no improvement in recall accuracy for the leading control details in the follow-up memory test compared to the initial memory test.

To summarize, inclusion of a non-suggested control group showed that the beneficial effects of suggestive interviewing are specific to the suggested details only. If the suggestions had promoted a general facilitation effect, then the three timing groups should have achieved higher accuracy scores on the non-suggested leading control details than the non-suggested control group; but this did not occur. Similarly, the three timing groups had the same scores on the misleading control details as the control group. This demonstrates that the misinformation that the three timing groups received did not negatively affect their memory for other details in the original event.

Summary of Experiment 1

The results showed that when the specific effects of timing are controlled by placing the interviews sufficiently distant from the event and the initial memory test, repeating suggestive interviews does not increase facilitation or misinformation effects. The results also show the long-term effect of suggestive interviewing, which was similar for all three timing conditions: First, there was a decrease in unprompted recall with time. Second, false recall increased on the total report measure, while the facilitation effects were the same in the 6-week initial memory test and the 27-week follow-up memory test. Third, although the relative size of the misinformation effect was smaller at the time of the

follow-up memory test compared to the initial memory test, this was due to an increase in false recall for the misleading control details and not a decrease in recall for the misleading suggestions; thus the smaller misinformation effects in the follow-up memory test may reflect *increased* distortion in the children's recall due to the inclusion of the misleading control details in the initial memory test rather than a diminished misinformation effect. Indeed, the children's reporting of the misleading control details in the follow-up memory test is another manifestation of a misinformation effect.

EXPERIMENT 2

Objectives and hypotheses

The results of Experiment 1 showed that the repetition of suggestions, when isolated from the specific influences of timing, had the same effect on the accuracy of children's recall as a single presentation of suggestions. Experiment 2 was designed to examine the relative effects of timing and repetition when the suggestive interviews are placed in a manner to *increase* consolidation effects (by placing the 'very early' suggestive interview 2 days after the event) or to *increase* misinformation effects (by placing the 'very late' suggestive interview 2 days before the memory test).

Experiment 2 extended the previous literature by including groups that received early or late suggestive interviews *and* a group receiving both an early and a late suggestive interview; in this way Experiment 2 fills in the cells that were missing in previous studies. For example, Warren and Lane (1995) and Gobbo (2000) argued that the presentation of misinformation soon after an event was the primary factor in decreasing children's recall accuracy for an experienced event. However, the comparisons made in these studies involved children who were re-exposed to the misinformation versus those who only received it at the early suggestive interview; neither of these studies included direct comparisons to children who had received misinformation one time only at a much later date. Such a design does not allow one to determine if between-group differences in memory result from the *timing* of the misinformation or from the *repetition* of the misinformation. In Experiment 2, it was hypothesized that if very early suggestive interviewing consolidated memory, then children who had a 'very early' suggestive interview would accurately recall more leading suggestions than children who only received the 'very late' suggestive interview. We also hypothesized that if the number of false reports was due to the strength of the misinformation trace, then children who received a 'very late' suggestive interview would demonstrate greater misinformation effects than children who only had a 'very early' suggestive interview. False reporting was expected to be highest for the group that received the misinformation twice because of repetition effects.

Method

Design

The design of Experiment 2 was identical to Experiment 1, with the following modifications: The total number of days between the magic show and the memory test was 39 instead of 40 ($M = 39$ days, $SD = 3.6$). The early suggestive interview occurred 2 days after the magic show ($M = 2$ days, $SD = 0.2$) and the late suggestive interview occurred 37 days following the magic show ($M = 36.8$ days, $SD = 3.6$)—that is, 2 days before the memory test; see Table 3. There was no 6-month follow-up memory test in Experiment 2.

Table 3. Mean number of days (with standard deviations) between the magic show and sessions in Experiments 1 and 2

	Experiment 1				Experiment 2			
	Timing condition				Timing condition			
Session	Mean delay	Early suggestive interview	Late suggestive interview	Two suggestive interviews	Mean delay	Very early suggestive interview	Very late suggestive interview	Two suggestive interviews
Suggestive interview 1	14.5 (1.6)	√		√	2.0 (0.2)	√		√
Suggestive interview 2	25.6 (2.6)		√	√	36.8 (3.6)		√	√
Initial memory test	40.7 (3.9)	√	√	√	39.0 (3.6)	√	√	√

As was the case in Experiment 1, children were randomly assigned to either the early suggestive interview condition, the late suggestive interview condition, or the two suggestive interviews condition. (For clarity, the early and late suggestive interviewing conditions in Experiment 2 are referred to as the 'very early' and 'very late' conditions.)

Participants

Ninety-six children attending kindergarten programmes were recruited from two schools in the Montreal area; these schools participated in Experiment 1, and thus the children were from similar neighbourhoods and similar middle-class socioeconomic backgrounds as the children in Experiment 1. The children were between the ages of 58 and 72 months. The mean age of the children (M age = 67 months) was similar to the mean ages of the children in Experiment 1 (M age = 68 months). All children had written parental permission to participate.

The children were randomly assigned to one of three timing conditions: the very early suggestive interview condition ($n = 31$), the very late suggestive interview condition ($n = 33$), or the two suggestive interviews condition ($n = 32$). The groups were equated in terms of age (the mean age of the children in each condition was 67 months) and approximately equated in terms of gender (42 to 52% of the children in the groups were girls).

Results

Analyses were conducted on free recall and total report of the magic show in the memory test. For all analyses, the between-group factor was timing condition (very early suggestive interview vs. very late suggestive interview vs. two suggestive interviews condition).

Event report

Free recall. A $3 \times 2 \times 2$ ANOVA with repeated measures was carried out on the number of utterances produced in free recall. Timing condition was the between-group factor and accuracy of utterance (accurate vs. inaccurate) and utterance status (suggested vs. non-suggested) as the repeated factors. There were significant main effects of accuracy,

Table 4. Mean number of free recall utterances (with standard deviations) in Experiment 2

Accuracy	Utterance status	Timing condition: Very early suggestive interview	Timing condition: Very late suggestive interview	Timing condition: Two suggestive interviews
Accurate	Suggested[a]	0.84 (0.69)	1.61 (1.03)	1.94 (1.13)
	Non-suggested	7.58 (4.21)	7.24 (4.53)	6.84 (3.55)
Inaccurate	Suggested[a]	0.19 (0.48)	0.58 (0.75)	1.13 (0.98)
	Non-suggested	0.87 (1.80)	1.24 (2.65)	1.28 (4.22)

[a]The maximum number of suggested details is 5.

$F(1, 93) = 157.22$, $p < 0.001$, and utterance status, $F(1, 93) = 134.11$, $p < 0.001$. There was also a significant interaction of Accuracy × Utterance Status, $F(1, 93) = 117.79$, $p < 0.001$. There were no significant main effects or interactions involving the timing condition variable. That is, timing condition did not influence the production rate or the accuracy rate of free recall. The data are presented in Table 4.

Overall, the children's free recall contained more accurate ($M = 4.34$, $SD = 4.15$) than inaccurate ($M = 0.89$, $SD = 2.34$) utterances. Planned comparisons performed on the Accuracy × Utterance Status interaction indicated that the *inaccurate* recall contained a similar number of suggested and non-suggested details (M suggested = 0.64, M non-suggested = 1.14), while the *accurate* recall contained more non-suggested details than suggested details (M suggested = 1.47, M non-suggested = 7.22), $p < 0.001$.

Total report. As in Experiment 1, total report accuracy was assessed by summing the accurate responses produced in free recall with the accurate responses on the recognition test.

The total number of accurate responses was analysed using a three-way repeated ANOVA with timing condition (very early suggestive interview vs. very late suggestive interview vs. two suggestive interviews) serving as the between-group factor and with accuracy (leading details vs. misleading details) and detail status (suggested details vs. control details) serving as the repeated factors. There were significant main effects of accuracy, $F(1, 93) = 71.03$, $p < 0.001$, and detail status, $F(1, 93) = 26.09$, $p < 0.001$. There was a significant two-way interaction of Accuracy × Detail Status, $F(1, 93) = 156.41$, $p < 0.001$. Importantly, there was a significant three-way interaction of Timing Condition × Accuracy × Detail Status, $F(2, 93) = 5.09$, $p < 0.01$. The data are presented in Table 5.

The three-way interaction was evaluated using planned comparisons. The facilitation effect was examined by comparing how each group performed on the leading suggested details compared to the leading control details. As hypothesized, the very late group did not receive the same benefit from receiving the suggestive interview as the very early and two interviews groups: Only the very early and the two interviews groups showed a significant facilitation effect, $p < 0.05$. Note that the facilitation effect was not due to decreased accuracy for the leading control details (the three groups performed similarly on these details: very early suggestive interview = 80%, very late suggestive interview = 75%, two suggestive interviews = 71%), but to increased accuracy on the leading suggested details (very early suggestive interview = 94%, very late suggestive interview = 85%, two suggestive interviews = 93%).

Table 5. Mean total report accuracy (with standard deviations) in Experiment 2

Accuracy	Detail status	Timing condition: Very early suggestive interview	Very late suggestive interview	Two suggestive interviews
Leading	Suggestion	4.71 (0.53)	4.24 (0.75)	4.66 (0.70)
	Control	4.00 (1.00)	3.78 (1.19)	3.53 (1.11)
Misleading	Suggestion	2.00 (1.46)	2.30 (1.70)	1.44 (1.37)
	Control	3.74 (1.13)	3.64 (1.43)	3.91 (1.40)

Note: The maximum for each cell is 5.

The misinformation effect was sizable and significant across all three timing groups; that is, the accuracy of rejecting misleading suggested details was lower than the accuracy for rejecting the misleading control details, $p < 0.001$. Planned comparisons supported the hypothesis that the two suggestive interviews group would show a greater misinformation effect than the very early and very late suggestive interview groups, $p < 0.05$.

Summary of Experiment 2

Very early leading suggestions resulted in a significant facilitation effect. Very early *or* very late misleading suggestions resulted in significant misinformation effects. Receiving both the very early and the very late suggestive interviews resulted in the facilitation effect *and* a heightened misinformation effect.

GENERAL DISCUSSION

Experiments 1 and 2 examined the relative influence of the timing and repetition of interviews containing leading and misleading suggestions on children's later recall accuracy. In contrast to previous experiments, our experiments included two design features that allowed a direct inspection of the independent and relative contributions of each of these factors. First, there were three timing conditions: children were interviewed either temporally close to the event (early), *or* temporally distant from the event (late), *or* both temporally close to the event and temporally distant from the event (early and late). Previous studies have only included two of these timing conditions. Second, we varied our definition of 'early' and 'late' so as to examine the separate and combined effects of consolidation and recency. In Experiment 1, we scheduled interviews so as to minimize these effects. In Experiment 2, we placed the interviews at times where it was reasonable to expect consolidation and recency effects separately and combined.

In Experiment 1, when the specific effects of timing were minimized, there were significant facilitation and misinformation effects regardless of the timing of the suggestive interviews and regardless of the number of suggestive interviews. In other words, repeated exposure to misinformation did not increase children's suggestibility and similarly under these same conditions repetition of true information did not increase facilitation effects.

Based on our review of the literature, the very early suggestive interview in Experiment 2 was timed to promote consolidation effects. In fact, we found that as

predicted, facilitation effects were obtained only for those children who received the very early suggestive interview. In summary, the results of Experiments 1 and 2 show that suggestive interviewing within 4 weeks of the event improved recall for the leading suggestions. The facilitation effect was not augmented by the addition of a second suggestive interview.

A different pattern occurred for misinformation. In Experiment 2, although all three timing groups (very early, very late, and two suggestive interviews) showed significant misinformation effects, children who received two suggestive interviews showed heightened misinformation effects compared to the children who had only one of the suggestive interviews. This is the first evidence that repeated presentation of misinformation increases children's suggestibility. Taken together, Experiments 1 and 2 suggest that timing moderates the effects of repetition: Misinformation effects were only heightened when the children experienced both the 'very early' and 'very late' suggestive interviews.

Our results are not consistent with those of Gobbo (2000) or Warren and Lane (1995), who failed to find increased suggestibility with repeated interviews. One explanation may be differences in the placement of the late interview in their studies and in our Experiment 2. We placed our suggestive interview 2 days before the memory test, whereas these other researchers placed the suggestions right before (in the same session) as the memory test. In the latter cases, the misinformation may have been so salient that it masked any cumulative effects of earlier suggestive interviews. It is interesting to note that the results of Experiment 2 are consistent with findings from Zaragoza and Mitchell (1996), who found that the repetition of misinformation increased adults' false reporting. In that study, testing occurred immediately, 2 days, or 1 week after exposure to the misinformation. At all retention intervals, subjects who had repeated exposures to the misinformation reported more false details than subjects who had only one exposure to the misinformation.

Experiment 1 also allowed us to investigate the long-term fate of facilitation and misinformation effects. Unlike Poole and Lindsay (2001), who found a significant decrease in misinformation effects over a 1-month period, the results of Experiment 1 showed remarkably little change between the initial memory test and the follow-up memory test about 21 weeks later. This may be related to the timing of the misinformation in the present study: In Poole and Lindsay's design, the misinformation was presented shortly before the initial memory test. In Experiment 1, the misinformation was presented 2 to 4 weeks before the initial memory test. As more-recently learned information is forgotten at a steeper rate than older information, it would be expected that the misinformation effects would drop off more quickly in Poole and Lindsay's study than in Experiment 1. A related explanation is that, because children's forgetting rates are steep, their forgetting for the event had reached asymptote by the 40-day initial memory test, and thus recall changed only very little over the next 5 months. These results indicate that misinformation effects can be particularly resistant to forgetting. It should be noted, however, that we did not include a group that received only the 6-month memory test, so our conclusions about the consequences of suggestive interviewing are based only on comparing groups that received some interpolated interviewing between the experience and the memory test.

The results of these studies are of practical importance in the interviewing of children. First, they suggest that even though suggestive interviewing may have ceased for a number of months, initial misinformation effects do not evaporate, particularly when children are asked direct questions about events. Second, although our results show the beneficial aspects of reminding children about true events, this may not be a simple strategy to

implement in the real world, where the interviewer may not know what events actually transpired. Consequently suggestions that are intended to facilitate children's long-term memory may be inaccurate and have the opposite result. Finally, although our results showed that repeated suggestive interviewing can be more detrimental to the accuracy of children's memory than a single suggestive interview, there still should be cause for concern even when only one suggestive interview has occurred.

ACKNOWLEDGEMENTS

This research was carried out at McGill University, Montreal, as part of LM's doctoral dissertation, supervised by MB, and supported by a Faculty of Graduate Studies and Research Fellowship from McGill University to the first author and a grant from the National Sciences and Engineering Research Council to the second author. We thank the children, their parents, and the staff of the participating schools for their involvement, interest, and enthusiasm. We also thank Carrie Marchionni, Anica Zarac, Amy Leach, Alexandra Fiocco, Christina Kyskan, Debbie Finkelberg, Debbie Mirman, Vicky Ferraro, and Joanne Simonetti for their assistance with collecting and coding the data. Finally, we thank Amye Warren and an anonymous reviewer for their helpful comments.

REFERENCES

Ackil, J. K., & Zaragoza, M. S. (1995). Developmental differences in eyewitness suggestibility and memory for source. *Journal of Experimental Child Psychology, 60*, 57–83.

Ackil, J. K., & Zaragoza, M. S. (1998). The memorial consequences of forced confabulation: age differences in susceptibility to false memories. *Developmental Psychology, 34*, 1358–1372.

Baker-Ward, L., Hess, T. M., & Flannagan, D. A. (1990). The effects of involvement on children's memory for events. *Cognitive Development, 5*, 55–69.

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.

Brainerd, C. J., & Poole, D. A. (1997). Long-term survival of children's false memories: a review. *Learning and Individual Differences, 9,* 125–152.

Bruck, M., Ceci, S. J., Francoeur, E., & Barr, R. J. (1995). I hardly cried when I got my shot!: Influencing children's reports about a visit to their pediatrician. *Child Development, 66*, 193–208.

Bruck, M., Ceci, S. J., & Francoeur, E. (2000). A comparison of three- and four-year-old children's use of anatomically detailed dolls to report genital touching in a medical examination. *Journal of Experimental Psychology: Applied, 6*, 74–83.

Bruck, M., Melnyk, L., & Ceci, S. J. (2000). Draw it again Sam: the effect of drawing on children's suggestibility and source monitoring ability. *Journal of Experimental Child Psychology, 77*, 169–196.

Bruck, M., Ceci, S. J., & Hembrooke, H. (2002). The nature of children's true and false narratives. *Developmental Review, 22*, 520–554.

Cassel, W. S., & Bjorklund, D. F. (1995). Developmental patterns of eyewitness memory and suggestibility: an ecologically based short-term longitudinal study. *Law and Human Behavior, 19*, 507–532.

Cassidy, D. J., & DeLoache, J. S. (1995). The effect of questioning on young children's memory for an event. *Cognitive Development, 10*, 109–130.

Ceci, S. J., & Bruck, M. (1995). *Jeopardy in the courtroom: A scientific analysis of children's testimony.* Washington DC: American Psychological Association.

Ceci, S. J., Ross, D. F., & Toglia, M. P. (1987). Suggestibility of children's memory: psycholegal implications. *Journal of Experimental Psychology, 116*, 38–49.

Ceci, S. J., Huffman, M. L. C., Smith, E., & Loftus, E. F. (1994). Repeatedly thinking about a non-event: source misattributions among preschoolers. *Consciousness and Cognition, 3*, 388–407.

Cohen, R. L., & Harnick, M. A. (1980). The susceptibility of child witnesses to suggestion. *Law & Human Behavior, 4*, 201–210.

Dent, H. R., & Stephenson, G. M. (1979). An experimental study of the effectiveness of different techniques of questioning child witnesses. *British Journal of Social and Clinical Psychology, 18*, 41–51.

Ebbinghaus, H. (1885/1913). *Memory: A contribution to experimental psychology* (H. A. Ruger, & C. E. Bussenius, Trans.). New York: Columbia University, Teacher's College. (Reprinted 1964, New York: Dover)

Fivush, R., & Hamond, N. R. (1989). Time and again: effects of repetition and retention interval on 2 year olds' event recall. *Journal of Experimental Child Psychology, 47*, 259–273.

Flin, R. (1995). Children's testimony: psychology on trial. In M. Zaragoza, J. R. Graham, G. C. N. Hall, R. Hirschman, & Y. S. Ben-Porath (Eds.), *Memory and testimony in the child witness* (pp. 240–254). Thousand Oaks CA: Sage.

Gobbo, C. (2000). Assessing the effects of misinformation on children's recall: how and when makes a difference. *Applied Cognitive Psychology, 14*, 163–182.

Goodman, G. S., & Reed, R. S. (1986). Age differences in eyewitness testimony. *Law and Human Behavior, 10*, 317–332.

Goodman, G. S., Taub, E. P., Jones, D. P. H., England, P., Port, L. K., Rudy, L., & Prado, L. (1992). Testifying in criminal court: emotional effects on child sexual assault victims. *Monographs of the Society for Research in Child Development, 57* (5, Serial No. 229).

Howe, M. L., Kelland, A., Bryant-Brown, L., & Clark, S. L. (1992). Measuring the development of children's amnesia and hypermnesia. In M. L. Howe, C. J. Brainerd, & V. F. Reyna (Eds.), *Development of long-term retention* (pp. 56–102). New York: Springer-Verlag.

Howe, M. L., O'Sullivan, J. T., & Marche, T. A. (1992). Toward a theory of the development of long-term retention. In M. L. Howe, C. J. Brainerd, & V. F. Reyna (Eds.), *Development of long-term retention* (pp. 245–255). New York: Springer-Verlag.

Hudson, J. A. (1990). Constructive processing in children's event memory. *Developmental Psychology, 26*, 180–187.

Huffman, M. L., Crossman, A. M., & Ceci, S. J. (1997). 'Are false memories permanent?': an investigation of the long-term effects of source misattributions. *Consciousness and Cognition, 6*, 482–490.

Leichtman, M. D., & Ceci, S. J. (1995). The effects of stereotypes and suggestions on preschoolers' reports. *Developmental Psychology, 31*, 568–578.

Loftus, E. F. (1975). Leading questions and the eyewitness report. *Cognitive Psychology, 7*, 560–572.

Marche, T. A. (1999). Memory strength affects reporting of misinformation. *Journal of Experimental Child Psychology, 73*, 45–71.

Marche, T. A., & Howe, M. L. (1995). Preschoolers report misinformation despite accurate memory. *Developmental Psychology, 31*, 554–567.

Memon, A., Wark, L., Bull, R., & Koehnken, G. (1997). Isolating the effects of the cognitive interview techniques. *British Journal of Psychology, 88*, 179–197.

Pezdek, K., & Roe, C. (1995). The effect of memory trace strength on suggestibility. *Journal of Experimental Child Psychology, 60*, 116–128.

Poole, D. A., & Lindsay, D. S. (2001). Children's eyewitness reports after exposure to misinformation from parents. *Journal of Experimental Psychology: Applied, 7*, 27–50.

Poole, D. A., & White, L. T. (1991). Effects of question repetition on the eyewitness testimony of children and adults. *Developmental Psychology, 27*, 975–986.

Powell, M. B., & Thomson, D. M. (1997). The effect of an intervening interview on children's ability to remember one occurrence of a repeated event. *Legal & Criminological Psychology, 2*, 247–262.

Roberts, K. P., Lamb, M. E., & Sternberg, K. J. (1999). Effects of the timing of postevent information on preschoolers' memories of an event. *Applied Cognitive Psychology, 13*, 541–559.

Tucker, A., Mertin, P., & Luszcz, M. (1990). The effect of a repeated interview on young children's eyewitness memory. *Australian & New Zealand Journal of Criminology, 23*, 117–124.

Warren, A. R., & Lane, P. (1995). The effects of timing and type of questioning on eyewitness accuracy and suggestibility. In M. Zaragoza, J. R. Graham, G. C. N. Hall, R. Hirschman, & Y. S. Ben-Porath (Eds.), *Memory and testimony in the child witness* (pp. 44–60). Thousand Oaks CA: Sage.

Zaragoza, M. S., & Mitchell, K. J. (1996). Repeated exposure to suggestion and the creation of false memories. *Psychological Science, 7*, 294–300.

APPENDIX

The Leading and Misleading Details in the Magician Event

Leading details	Misleading details
Scene 1: Setting up	
Magician wore a black hat	Magician wore black gloves
Child wore a red magic helper cape	Child wore red magic helper boots
Magician told child a magic word	Magician blew up a balloon
Magician took a toy rabbit out of a box	Magician took a toy rabbit out of her sleeve
Scene 2: Ball trick	
Magician put the red ball in her pocket	Magician put the red ball under the table
Magician waved the wand	Magician dropped the wand
Child honked the horn	Child rang a bell
Magician covered her hand with a scarf	Magician covered her hand with some paper
Scene 3: Water trick	
Child poured water into the cup	Magician drank the water in the cup
Water spilled in the bowl	Water spilled on the floor
Magician stirred the water with a spoon	Magician stirred the water with her finger
Magician put the cup in the corner	Magician put the cup in the garbage
Scene 4: The fall	
Magician fell on her shoelaces	Magician fell on a toy
Magician laughed when she fell	Magician cried when she fell
Magician hurt her hand	Magician hurt her leg
Magician put some band-aids on herself	Child put some band-aids on the magician
Scene 5: Clean-up	
Magician put the tricks away in a box	Child put the tricks away in a box
Magician brushed her stage with a little brush	Magician brushed her hair with a little brush
Magician showed the child a picture of her puppy	Magician sang the child a song
Magician gave the child an award for being a good helper	Magician gave the child a hug for being a good helper

