Eyewitness confidence : the relation between accuracy and confidence in episodic memory

Odinot, G.

4

Repeated partial eyewitness questioning causes confidence inflation but not retrieval-induced forgetting*

* A short version of this chapter is in press and published online: Odinot, G., Wolters, G., & Lavender, T. (in press). Repeated partial eyewitness questioning causes confidence inflation but not retrieval-induced forgetting. *Applied Cognitive Psychology*. http://dx.doi.org/10.1002/acp.1443

Summary

During a crime investigation eyewitnesses are often interviewed more than once. Repeated post-event questioning offers an opportunity for retrieval practice. Practicing retrieval of a subset of memories may suppress access to related memories, a phenomenon known as retrieval-induced forgetting. In this paper we investigated the generalization of retrieval-induced forgetting to episodic eyewitness memory of a complex event. The results indicated that repeated retrieval improves future recall of practiced information, but does not induce forgetting of related information. Retrieval practice resulted in higher confidence ratings, both for correct and incorrect answers. The practical consequence of this latter finding is that repeated questioning should be avoided if possible, because it may lead to artificially high confidence levels.

Introduction

Repeated retrieval of memory traces can have consequences for confidence and the amount of information retrieved during later recall attempts. Typically, prior retrieval of information increases the probability of the same information being retrieved at a later recall attempt. Occasionally, also information may be retrieved that was not recalled previously, a phenomenon known as spontaneous recovery or hypermnesia (e.g., Roediger, McDermott, & Goff, 1997; Scrivner & Safer, 1988; Turtle & Yuille, 1994). Repeated retrieval of particular information, however, may also lead to diminished accessibility of other related information that was not retrieved in prior recall attempts. This phenomenon is called retrieval-induced forgetting (Anderson, Bjork, & Bjork, 1994; Anderson & McCulloch, 1999; Anderson & Spellman, 1995; Barnier, Hung, & Conway, 2004; MacLeod, 2002; Shaw, Bjork, & Handal, 1995)

The positive and negative effects of repeated retrieval or retrieval practice are highly relevant to the study of eyewitness reports. During the investigation of a crime, eyewitnesses may be asked to provide a description of the event. This initial interview is often followed by additional interviews during later stages of the investigation. One of the reasons to question witnesses several times is that witnesses may provide new information during follow-up questioning. Information that could not be remembered initially may be remembered later (Penrod & Cutler, 1995). However, repeated interviewing may also introduce distortion of memory, and it offers witnesses the opportunity to practice retrieval of their memories. Research suggests that such retrieval practice indeed may affect the amount of information recalled, as well as the level of confidence that is expressed by witnesses about the accuracy of their memory (e.g., Granhag, 1997; Roediger et al., 1997; Shaw, 1996; Shaw et al., 1995; Shaw & McClure, 1996).

Retrieval-induced forgetting was reported first in studies showing that recall of unpracticed exemplars of categories was impaired by practicing recall of other exemplars. This impair-

ment occurs not only relative to recall of practiced exemplars, but also relative to a baseline of recall of exemplars from other unpracticed categories. These memory failures are attributed to an inhibition mechanism that might arise from the executive control processes that resolve interference between competing memory traces (see for more detailed discussions, e.g., Anderson et al., 1994; Anderson & Spellman, 1995; Levy & Anderson, 2002; MacLeod, 2002; MacLeod & Macrae, 2001).

Originally it was argued that retrieval-induced forgetting occurs because the strong semantic links in our memory required an active suppression of the non-practiced items, which resulted in robust effects of retrieval-induced forgetting in laboratory studies. However, retrieval-induced forgetting may also occur in the absence of pre-existing category-exemplar relationships. Such findings were reported for instance by Ciranni & Shimamura (1999) and Gomez-Ariza, Lechuga, Pelegrina, & Bajo (2005). Both studies investigated retrieval-induced forgetting for newly learned (episodic) associations. They showed that also newly formed associations between perceptual features, such as shape and color, are vulnerable to the inhibitory effects of retrieval practice, thus extending the domain of retrieval-induced forgetting from semantic memory to episodic memory.

Additional studies have shown that retrieval-induced forgetting also generalizes to conditions that more resemble real-life situations. For instance, Macrae & MacLeod, (1999) showed that retrieval-induced forgetting even occurs when explicit instructions to remember the stimulus material are absent. This is exactly what happens to real-world witnesses, as they do not know in advance the importance of the events that they perceive. And Barnier et al., (2004) reported retrieval-induced forgetting effects for autobiographical memories.

Shaw & McClure, (1996) tested the effect of retrieval-induced forgetting in eyewitness memory using more complex stimulus material than the simple word lists mostly used in research in this paradigm. In an initial phase, participants were shown color slides of the interior of a student's apartment and they were asked to memorize the objects shown. The slides contained two categories of objects (college textbooks and college sweatshirts). After viewing the slides the participants were questioned three times over a 20 minutes interval about half of the objects from one of the two categories. Following this repeated retrieval phase, participants were asked to recall as many of the target items as they could. Recall of practiced items was higher than that of unpracticed items from both the practiced and unpracticed categories, demonstrating a retrieval practice effect. More importantly, recall of unpracticed items of the practiced category was lower than that of unpracticed items of the unpracticed category, indicating that retrieval-induced forgetting indeed occurred. An interesting feature of this study was that also a control group was added that was not questioned at all prior to the final test. Recall in this control group appeared to be lower than recall of practiced items, but higher than recall of unpracticed items. In a similar vein, MacLeod (2002) reported that retrieval-induced forgetting can occur for the recall of details of the description of a suspect following repeated questioning on a subset of these details.

These findings have important practical implications because they suggest that repeatedly questioning witnesses can actually lead to poorer recall of previously neglected details. According to MacLeod (2002) the likelihood that retrieval-induced forgetting occurs in real-life settings, like interviewing eyewitnesses, may be quite high. Interviews of eyewitnesses often constitute incomplete retrieval tasks as police officers and other investigators tend to limit their questions to specific aspects of the incident (e.g., the burglar's weapon, see Shaw et al., 1995). Thereby, retrieval practice on a subset of the total memory might have an inhibiting effect on later recall of related information that was not part of the initial questioning. Details that became important during the course of an investigation may be difficult to remember because they are inhibited by information retrieved in earlier interviews that focused on other details.

Retrieval practice has been shown to prompt forgetting of related information in an increasing number of studies, and several authors have warned already for the potential problems for eyewitness reliability. However, the problem may not be as large as suggested, because also boundary conditions for the occurrence of retrieval-induced forgetting have been reported (see for an overview Levy & Anderson, 2002). Such boundary conditions may render the occurrence of retrieval-induced forgetting in real life eyewitness settings less likely. One of the boundary conditions is that it probably is a relatively short-lived phenomenon. MacLeod & Macrae (2001; see also Macrae & MacLeod, 1999) argued that retrieval-induced forgetting can be seen as an adaptive mechanism. By actively suppressing competing memories that share the same retrieval cue as the target memory, people can ensure that they only become aware of the recollection that is relevant for their current cognitive goals. This inhibition, however, should be relatively short-lived in order for it to be socially adaptive. Once people have achieved their current cognitive goals, the inhibitory effect should cease to operate.

The level of integration between competing memories has been suggested as another boundary condition. Anderson & McCulloch (1999) reported that instructions for participants to form interconnections between category exemplars reduced retrieval-induced forgetting. Highly integrated episodic memories therefore may be less susceptible to the disruptive effects of retrieval practice than poorly integrated memories.

In sum, the likelihood that retrieval-induced forgetting occurs in real life seems to be to quite high. Several studies indeed have shown robust retrieval-induced forgetting effects in eyewitness related settings. On the other hand, however, the effect of retrieval-induced forgetting is restricted by boundary conditions that make the occurrence of retrieval-induced forgetting in real life eyewitness situations less obvious. So far, no study has reported retrieval-induced forgetting in a complex dynamic event. It remains to be seen therefore whether retrieval-induced forgetting will occur in eyewitness memory of a complex episodic event and with relatively long retention intervals. This is the main purpose of this study.

In addition, we will look at evidence for the opposite phenomenon of hypermnesia; i.e., remembering information on a later occasion that could not be remembered in previous recall attempts. The possibility that later retrieval attempts may produce new additional information is an important argument for repeated questioning of eyewitnesses. Therefore, we will also investigate if this presumed positive result of repeated questioning compensates for possible negative effects of retrieval-induced forgetting.

Effects of retrieval practice on confidence

Experimental evidence suggests that retrieval practice does not only influence later recall in terms of the amount of information retrieved, but also in terms of the confidence people have in their memory reports. Shaw (1996) for example, conducted a series of experiments to investigate the effect of repeated post-event questioning on eyewitness confidence. He reported that repeated post-event questioning did not increase accuracy, but participants expressed greater confidence in their repeated responses, irrespective of whether responses were accurate or inaccurate. According to Shaw, Bjork and Handal (1995), such findings can be explained by assuming that repeated retrieval of an episodic memory leads to an increase in 'retrieval fluency' for that specific memory item and greater ease of future retrieval. Ease of retrieval has been shown to serve as a basis for confidence judgments (Kelley & Lindsay, 1993). The retrieval-fluency hypothesis can explain why both correct and incorrect memories suffer from confidence inflation due to repeated post-event questioning (Shaw & McClure, 1996).

Confidence inflation without a corresponding increase in accuracy is problematic in a legal context. There is a widely held intuitive belief that confidence expressed about a memory can be used to infer its accuracy, both in the general public as well as by legal professionals (Cutler, Penrod, & Stuve, 1988; Leippe, 1980; Lindsay, Wells, & O'Connor, 1989; Luus & Wells, 1994; Penrod & Cutler, 1995; Wise & Safer, 2004). The relationship between confidence and accuracy depends on numerous factors, like kind of tests used (Robinson & Johnson, 1996), kind of material presented (Perfect, Watson, & Wagstaff, 1993; Roediger & McDermott, 1995), distribution of item difficulty (Kebbell, Wagstaff, & Covey, 1996), and personality characteristics (Bothwell, Brigham, & Pigott, 1987; Nolan & Markham, 1998). However, the general finding is that the relationship between accuracy and confidence is far from perfect. Therefore, it is obvious that any spurious inflation of confidence without a corresponding increase in accuracy is potentially harmful in a judicial context.

The goal of the present study is to investigate whether retrieval-induced forgetting, hypermnesia and confidence inflation generalize to highly complex stimuli and a time span that is more realistic for real-life eyewitness situations. Most studies addressing these questions used simple word lists or series of static slides as stimuli instead of complex dynamic events unfolding in a meaningful context like in a real life situation. Furthermore, the forced-choice recognition tasks that are often used in eyewitness research bear little resemblance

to the free and cued recall questions asked in actual police interviews. Also time intervals between retrieval practice and final testing, especially in experiments studying retrieval-induced forgetting, are relatively short compared to what witnesses experience in real life. Therefore, in this experiment a 20 minute video is used as stimulus material during the study phase. This is followed by two retrieval practice sessions and a final test that are separated by intervals of several weeks. Participants are given cued recall tasks in which one half of the questions is repeatedly presented, and the other half is presented during final testing only. In these respects the design in this study differs substantially from previous research. Our aim is to see whether retrieval-induced forgetting occurs under these conditions. This issue is important because retrieval-induced forgetting could hamper criminal investigations. Confidence inflation for incorrect responses to practiced questions also poses a potential problem, because legal professionals often rely intuitively on confidence as an indicator of witness accuracy.

Method

Participants

Sixty-three students, 50 female and 10 male, participated in the experiment, either for credits or a financial reward. Their age ranged from 18 to 39 years, with a mean of 22 years.

Materials and procedure

Videotape. Participants were individually shown a television programme previously broadcasted on Dutch television. The video was converted into a MPEG-file which was shown individually to the participants on a 17 inch monitor. The video depicts two different storylines that leads to an accident between two cars. The duration of the video was 20 min. Participants were not explicitly instructed to memorize the contents of the video but were aware participating an eyewitness experiment. None of the participants indicated that they had seen the video before.

Questionnaires. In the final test session, five weeks after viewing the video, all participants filled out a 30-item open-ended questionnaire about details in the video. For half of the participants, this final test session was preceded by two retrieval practice sessions (after one and three weeks) in which they had to answer half the questions of the complete list. Each questionnaire started with a general question wherein the participants were asked to describe the two story lines in global and general terms. As is common in police interviews, this general question is asked in order to reinstate and refresh the memory before proceeding with more specific questions. Next, participants answered in writing a series of questions that varied in specificity from relatively global (e.g., "Can you describe how the two cars ran into each other?") to highly detailed (e.g., "What was the colour of the car that

the two brothers were planning to buy?"). It was stressed that answers should be as complete and detailed as possible, corresponding with their memory for details or scenes from the video. However, it was also stressed that if participants could not remember the answer they should refrain from answering by indicating "do not know".

To allow a fine-grained analysis of the responses, participants were instructed to give their answers in the form of small units of information. A unit was described as a single element or aspect of information. To explain this to participants, the following example was given. *Question:* 'What did the boy do after he was refused to enter a nightclub?'; *answer:* 'he went home'; 'on his green bicycle'; 'he took a silver coloured pistol'; 'out of the top drawer'; 'of his bedside table'; 'he went back to the nightclub'; 'were he shot the doorkeeper'. To encourage the subjects to give single units of information to the more global questions, the lines on the answering sheet were restricted in length. Participants could answer with as many units of information as they needed.

The complete 30-question list was composed of two 15 question blocks, labelled A and B. In each block, every question corresponded to a question in the other block by asking about a different aspect of the same scene from the video. In this way we created potential recall competition to enhance the possible occurrence of retrieval-induced forgetting. For example, the question "What was the colour of the car that the two brothers were planning to buy?" (block A) corresponded with "What was the colour of the car that the two brothers took for a test ride?" (block B). These details were shown briefly after one another in the same video-scene. The two blocks of the 15 questions were matched in terms of difficulty on the basis of the results from a pilot study.

During the final test session, participants were instructed to make a confidence judgment for each unit of information they provided. They were asked to indicate their confidence regarding the accuracy of each unit of information given on a 7-point Likert scale ranging from 1 (very uncertain) to 7 (absolutely certain). This was asked in the final session only, in order to avoid possible effects of repeated confidence judgments, such as remembering earlier ratings and being motivated to appear consistent.

Design

Retrieval practice was manipulated between participants. Half of the participants received retrieval practice on a subset of their memories (either questions of block A or B), on two occasions; 1 and 3 weeks after seeing the video. So for each question that was practiced, the corresponding question in the other block was not practiced. The other half of the participants did not receive any retrieval practice. For all participants the final test session took place 5 weeks after the video. During the final test session, all participants filled out the complete 30-item questionnaire; each practiced question was directly followed by the corresponding but non-practiced question. This design allows to compare recall of practiced questions (Rp+) with recall of related but unpractised questions (Rp-) within subjects, and

to compare between subjects recall of practised/unpractised questions with recall of the same questions without any retrieval practice (Nrp).

Scoring. All units of information recalled by the participants were scored as correct or incorrect. Information was scored *correct* when it matched the information in the video. *Incorrect* information consists of units of information not presented in the video, information that was either incorrectly remembered or fantasized by the participants. Two experimenters did the scoring, and in case of a disagreement, a third experimenter settled the dispute. Of all units generated less than 0.5% could not be classified as correct of incorrect; these were discarded from further analysis.

Results

After removal of one outliers with very poor accuracy (11.4% correct), the data from the remaining sixty-two participants were analyzed. Per subject, the number correct units of information, the number incorrect units of information and the number of "do not know" answers were determined. Overall the average number of "do not know" answers per block of 15 questions in the final test was 2.1 (14 %). Although the average number of "do not know" answers was somewhat lower in the repeated retrieval condition (1.5) than in the unpractised (2.1) and in the no practice at all conditions (2.5), these differences were not significant. Because the number of correctly and incorrectly recalled units of information differs over questions and participants, we will present and analyse these data as the actual (average) numbers per block of 15 questions.

Retrieval practice – correct recall. The mean numbers of correctly and incorrectly recalled units of information per condition (i.e., averaged over blocks of 15 questions) are shown in Table 1. First, we assessed the effect of retrieval practice on correctly recalled units of information by comparing recall of the practiced (Rp+) and the unpractised but related question (Rp-) lists. A paired samples t-test showed a significant difference in recall between practiced (M = 22.5) and unpractised (M = 15.7) questions, t (28) = 5.19, p < 0.001. The effect of retrieval practice was confirmed by comparing recall of practiced questions (Rp+) with recall of the control group that did not receive practice at all (Nrp). An independent samples t-test showed a significant difference (t (60) = 4.36, p < .001). These results clearly show a positive effect of retrieval practice; it reduces forgetting with longer retention intervals.

To determine if retrieval-induced forgetting occurred, we compared recall of unpractised but related (Rp-) questions with recall in the control group (Nrp). This difference was not significant (t (60) = 0.33, p = n.s.). Actually recall of the Rp- lists was slightly better (M = 15.7) than recall in the control group (M = 15.2). This finding indicates that in the conditions studied here, there is no evidence for retrieval-induced forgetting.

We also looked at the possible occurrence of hypermnesia in correctly recalled information units, i.e., correct recall of elements in later retrieval attempts that were not recalled in previous attempts. To determine the occurrence of hypermnesia, we counted the number of correctly recalled units of information that were absent in a previous recall attempt. Overall, we found only 49 instances of hypermnesia (an average of 1.6 per subject), 30 in the 1-3 weeks comparison, 19 in the 1-5 weeks comparison and no new correct items in the 3-5 weeks comparison.

Table 1 Mean number of correctly recalled units of information per condition (i.e., blocks of 15 questions, the mean of the no retrieval control group is averaged over 2 blocks). Standard deviations in parentheses.

	Mean number of correct units of information			
	1 week	**3 weeks**	**5 weeks**	
	Rp+	**Rp+**	**Rp+**	**Rp-**
Retrieval practice N=29	26.1 (10.0)	24.2 (7.6)	22.5 (7.5)	15.7 (4.5)
			Nrp	
No retrieval practice N=33	-	-	15.2 (5.4)	

Retrieval practice – incorrect recall. To test the possibility that retrieval practice may lead to an increase in errors, we analyzed recall errors in the same way as correct recall. The average numbers of incorrect units of information in each condition are presented in Table 2. Although there seem to be slightly more errors in the repeated retrieval conditions, none of these analyses showed any significant difference related to the retrieval practice manipulations (all *t* values < 1.0).

Table 2 Mean number of incorrectly recalled units of information per condition (i.e., blocks of 15 questions). Standard deviations in parentheses.

	Mean number of incorrect units of information			
	1 week	**3 weeks**	**5 weeks**	
	Rp+	**Rp+**	**Rp+**	**Rp-**
Retrieval practice N= 29	7.5 (4.2)	7.4 (3.7)	7.8 (4.7)	6.8 (2.6)
			Nrp	
No retrieval practice N=33	-	-	6.9 (3.2)	

Confidence inflation. Confidence ratings on a 7-point scale (from 1 = very uncertain to 7 = absolutely certain) were requested during the final test only. Mean confidence ratings were determined for correctly and incorrectly recalled information units in all conditions, and

these are shown in Table 3. Overall, confidence was significantly higher for correct units of information (M = 5.4) than for incorrect units (M = 4.5), t (61) = 11.2, p < .001.

After retrieval practice, participants showed more confidence in their correct answers to practiced questions (M = 5.9) than in their correct answers to unpractised questions (M = 5.4), t (28) = 4.84, p < 0.001. Comparisons with confidence in accurate recall of the control group that did not receive retrieval practice (M = 5.2), showed a significant difference with practiced questions (t (60) = 3.88, p < .001), but no difference with unpractised questions (t (60) = 1.01, p = n.s.).

Analysis of confidence in incorrect answers yielded the same pattern of results as for correct answers. Participants who had received retrieval practice showed more confidence in their incorrect answers to practiced questions (M = 4.9) than in their incorrect answers to unpractised questions (M = 4.4), t (28) = 2.1, p < 0.05. Comparisons with the control group again showed a significant difference with practiced questions (t (60) = 2.83, p < .01), but no difference with unpractised questions (t (60) = .92, p = n.s.). The analyses of the confidence ratings provide converging evidence that indeed retrieval practice does result in confidence inflation, both for correct and incorrect answers.

Table 3 Mean confidence ratings for correct and incorrect units of information in the retrieval practice conditions (scale values: 1=very uncertain to 7=absolutely certain). Standard deviations between parentheses.

	Mean confidence in final recall session			
	Correct		**Incorrect**	
	Rp+	**Rp+**	**Rp+**	**Rp-**
Retrieval practice N= 29	5.9 (.62)	5.4 (.83)	4.9 (1.02)	4.4 (.99)
	Nrp		**Nrp**	
No retrieval practice N= 33	5.2 (.82)		4.2 (1.09)	

Confidence – accuracy relationship. To analyze the accuracy-confidence relation, we determined the number of correct and incorrect units of information for each confidence level over all participants, and we calculated Goodman-Kruskal gamma correlation coefficients. The correlation between accuracy and confidence in the group with no retrieval practice was G = 0.41. In the retrieval practice group, the Gamma correlation for the Rp+ items was G = 0.45, and for the Rp- items G = 0.29. All these correlations are highly significant (all p < 0.001). To gain more insight in the confidence-accuracy relationship, we calculated the total number and the proportion of correct units of information as a function of confidence levels. As can be seen in Table 4, five weeks after seeing the video, answers given a confidence judgement at the lower end of the scale were as often correct as incorrect. Of the answers judged with the highest confidence score, 84% was correct. This shows that information recalled with the highest confidence has a high probability of being accurate, although even here chances of erroneous recall are not negligible.

Table 4 Distribution of units of information along the confidence scale.

	Confidence scale							
	1	**2**	**3**	**4**	**5**	**6**	**7**	**Total**
# Answers	129	214	306	343	478	541	1056	3067
Prop. correct	0.50	0.49	0.61	0.57	0.67	0.75	0.84	0.71

Discussion

The main goal of the present study was to investigate whether retrieval-induced forgetting generalizes to real-life eyewitness situations. In addition, we looked at hypermnesia and confidence inflation. To create a more ecologically valid situation, we used complex stimuli, a cued recall format for repeated questioning, and retention intervals of several weeks. Under these conditions we found no evidence for retrieval-induced forgetting. Consistent with previously reported results, retrieval practice did have a positive effect on recall of practiced information, and it caused confidence inflation; participants became more confident, both in correct and incorrect answers. Hypermnesia did occur, but only to a very limited extent.

There are several explanations for the absence of retrieval-induced forgetting in the conditions as used in this study. First, the structure of the event memory created by our stimulus material obviously is much more complex than the hierarchical structure of word and picture categories that is often used in studies on retrieval-induced forgetting. It is likely that the memories from the video were highly integrated because they formed a meaningful, coherent narrative story. Interconnections between episodic memories ensure multiple retrieval pathways to find and give access to information. This memory integration may serve as a safeguard against retrieval-induced forgetting, as was shown by Anderson & McCulloch (1999). Highly integrated episodic memories, such as a coherent video narrative as used in this experiment, or a real-life event, are probably less vulnerable to the inhibitory effects of retrieval practice than recollections of isolated words and pictures which are organized by their categorical similarity.

Second, the recall procedure with item-specific cues could be a reason why retrieval-induced forgetting did not occur. According to MacLeod and Macrae (2001), retrieval-induced forgetting serves to suppress competing memories that share the same retrieval cue as a target memory. In our study, however, the overlap of retrieval cues for accessing a specific item in a highly integrated episodic memory was only partial. Although each pair of corresponding questions referred to the same scene in the video, different elements had to be recalled and the overlap of the retrieval cues was not complete. Therefore, recall competition may have been relatively small, reducing the need for retrieval inhibition.

A third explanation for the absence of retrieval-induced forgetting might be the retention interval between stimuli and recall sessions. Final testing took place 5 weeks after the

study phase and two weeks after the second retrieval practice session. This exceeds by far the intervals that are typically used in studying retrieval-induced forgetting. Although MacLeod and Macrae (2001) still found some retrieval-induced forgetting after a 24 hour interval, they concluded that it probably is a relatively short-lived phenomenon. They argued that once the target memory has been retrieved, it is no longer necessary to block access to competing memories. Whichever explanation is correct, our results indicate that retrieval-induced forgetting either does not occur in the conditions used, or does not persist over repeated retrieval intervals of several weeks.

Our findings showed a significant inflation of witness confidence for the retrieval practice items, regardless whether the answers were correct or incorrect. This seems to be in conflict with results showing some, but not significant, confidence inflation with repeated recall (Odinot & Wolters, 2006). In that study, however, confidence levels were very high which probably precluded confidence inflation to show up because of a ceiling effect. The present result does replicate findings of Shaw and McClure (Shaw, 1996; Shaw & McClure, 1996) who also reported that repeated post-event questioning results in higher confidence levels. According to the 'retrieval fluency' hypothesis, post-event questioning about an episodic memory leads to strengthening and consequently an increase in retrieval fluency for the information that is recalled, which in turn may result in elevated levels of confidence. Because repeated post-event questioning affects retrieval fluency independent of the correctness of the response, confidence inflation occurs both for correct and incorrect memories.

The confidence-accuracy relations found in this study appeared to be independent of presence or absence of retrieval practice. The relationship suggests that confidence may be a moderately useful indicator for accuracy, because selecting only answers with the highest confidence rating results in filtering out a large proportion of incorrect information. Unfortunately, however, even with the highest confidence ratings there still remains a substantial proportion of incorrect information. Therefore, no single witness statement can be accepted as certainly correct on the basis of confidence alone.

We believe the result of this study have three important practical implications. First, witnesses should be questioned as soon as possible after an event, because delays reduce the amount of information that can be recalled. This, of course, is not a new finding, but it again emphasizes the importance of questioning witnesses as soon as possible after an event. Second, it seems unlikely that repeated questioning in real life conditions may cause retrieval-induced forgetting. Prior studies already suggested limiting factors that would make retrieval-induced forgetting less likely in everyday conditions, and the results of the present study bear evidence to this conclusion. We acknowledge, however, that such forgetting can not be completely ruled out in other conditions. Third, although repeated questioning probably does not cause retrieval-induced forgetting, it nevertheless should be avoided as much as possible. Our findings do not show strong indications that repeated

recall attempts enhance memory (i.e., hypermnesia), but they do indicate that it does cause confidence inflation which is a potential problem in a legal context. Since confidence inflation occurs both for correct and incorrect answers, the tendency to rely on confidence as an indicator for accuracy enhances the probability that incorrect information is falsely accepted.

A number of questions remain regarding retrieval-induced forgetting in eyewitness memory. Several authors have suggested that retrieval-induced forgetting is highly likely to occur in some real life eyewitness situations (e.g., McLeod, 2002; Shaw et al. 1995). For instance, intensive repeated partial questioning of suspects is common practice in crime investigations. The Cognitive Interview (Fisher & Geiselman, 1992), a memory enhancing technique to help eyewitnesses to remember as much as possible during a police interview, involves deliberate repeated recall of the memory. How such repeated retrieval practices in a relative short time interval influence memory is still unclear. If retrieval-induced forgetting would occur in these situations, the question arises for how long this inhibition exists, and whether there is even a chance that the unpracticed items become unavailable permanently. Clearly, further research is needed to see what the specific conditions are for retrieval-induced forgetting to occur in real life eyewitness situations.

References

Anderson, M. C., Bjork, R. A., & Bjork, E. A. (1994). Remembering can cause forgetting: Retrieval dynamics in long-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*, 1063-1087.

Anderson, M. C., & McCulloch, K. C. (1999). Integration as a general boundary condition on retrieval-induced forgetting. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25*, 608-629.

Anderson, M. C., & Spellman, B. A. (1995). On the status of inhibitory mechanisms in cognition: Memory retrieval as a model case. *Psychological Review, 102*, 68-100.

Barnier, A. J., Hung, L., & Conway, M. A. (2004). Retrieval- induced forgetting of emotional and unemotional autobiographical memories. *Cognition and emotion, 18*, 457-477.

Bothwell, R. K., Brigham, J. C., & Pigott, M. A. (1987). An exploratory study of personality differences in eyewitness memory. *Journal of Social Behavior and Personality, 2*, 335-343.

Ciranni, M. A., & Shimamura, A. P. (1999). Retrieval-induced forgetting in episodic memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25*, 1403-1414.

Cutler, B. L., Penrod, S. D., & Stuve, T. E. (1988). Juror decision making in eyewitness identification cases. *Law and Human Behavior, 12*, 41-55.

Fisher, R. P., & Geiselman, E. R. (1992). *Memory-enhancing techniques for investigative interviewing. The cognitive interview*. Springfield, Illinois: Charles C. Thomas.

Gomez-Ariza, C. J., Lechuga, M. T., Pelegrina, S., & Bajo, M. T. (2005). Retrieval-induced forgetting in recall and recognition of thematically related and unrelated sentences. *Memory & Cognition, 33*, 1431-1441.

Granhag, P. A. (1997). Realism in eyewitness confidence as a function of type of event witnessed and repeated recall. *Journal of Applied Psychology, 82*, 599-613.

Kebbell, M. R., Wagstaff, G. F., & Covey, J. A. (1996). The influence of item difficulty on the relationship between eyewitness confidence and accuracy. *British Journal of Psychology, 87*, 653-662.
Kelley, C. M., & Lindsay, D. S. (1993). Remembering mistaken for knowing: Ease of retrieval as a basic for confidence in answers to general knowledge questions. *Journal of Memory and Language, 32*, 1-24.
Leippe, M. R. (1980). Effects of integrative memorial and cognitive processes on the correspondence of eyewitness accuracy and confidence. *Law and Human Behavior, 4*, 261-274.
Levy, B. J., & Anderson, M. C. (2002). Inhibitory processes and the control of memory retrieval. *Trends in cognitive sciences, 6*, 299-305.
Lindsay, R. C. L., Wells, G. L., & O'Connor, F. J. (1989). Mock-juror belief of accurate and inaccurate eyewitnesses. *Law and Human Behavior, 13*, 333-339.
Luus, E. C. A., & Wells, G. L. (1994). Eyewitness identification confidence. In D. F. Ross & J. D. Read & M. P. Toglia (Eds.), *Adult Eyewitness Testimony* (pp. 348-361). Cambridge: Cambridge University Press.
MacLeod, M. D. (2002). Retrieval-induced forgetting in eyewitness memory: Forgetting as a consequence of remembering. *Applied Cognitive Psychology, 16*, 135-149.
MacLeod, M. D., & Macrae, C. N. (2001). Gone but not forgotten: The transient nature of retrieval-induced forgetting. *Psychological Science, 12*, 148-152.
Macrae, C. N., & MacLeod, M. D. (1999). On recollections lost: When practice makes imperfect. *Journal of Personality and Social Psychology, 77*, 463-473.
Nolan, J., & Markham, R. (1998). The accuracy-confidence relationship in an eyewitness task: Anxiety as a modifier. *Applied Cognitive Psychology, 12*, 43-54.
Odinot, G., & Wolters, G. (2006). Repeated recall, retention interval and the accuracy-confidence relation in eyewitness memory. *Applied Cognitive Psychology, 20*, 973-985.
Penrod, S. D., & Cutler, B. L. (1995). Witness confidence and witness accuracy: Assessing their forensic relation. *Psychology, Public Policy, and Law, 1*, 817-845.
Perfect, T. J., Watson, E. L., & Wagstaff, G. F. (1993). Accuracy of confidence ratings associated with general knowledge and eyewitness memory. *Journal of Applied Psychology, 78*, 144-147.
Robinson, M. D., & Johnson, J. T. (1996). Recall memory, recognition memory, and the eyewitness confidence-accuracy correlation. *Journal of Applied Psychology, 81*, 587-594.
Roediger, H. L., & McDermott, K. B. (1995). Creating false memories: remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*, 803-814.
Roediger, H. L., McDermott, K. B., & Goff, L. M. (1997). Recovery of true and false memories: Paradoxical effects of repeated testing. In M. A. Conway (Ed.), *Recovered memories and false memories* (pp. 118-149). Oxford: Oxford University Press.
Scrivner, E., & Safer, M. A. (1988). Eyewitnesses show hypermnesia for details about a violent event. *Journal of Applied Psychology, 73*, 371-377.
Shaw, J. S. (1996). Increases in eyewitness confidence resulting from postevent questioning. *Journal of Experimental Psychology: Applied, 2*, 126-146.
Shaw, J. S., Bjork, R. A., & Handal, A. (1995). Retrieval-induced forgetting in an eyewitness-memory paradigm. *Psychonomic Bulletin & Review, 2*, 249-253.
Shaw, J. S., & McClure, K. A. (1996). Repeated postevent questioning can lead to elevated levels of eyewitness confidence. *Law and Human Behavior, 20*, 629-653.
Turtle, J. W., & Yuille, J. C. (1994). Lost but not forgotten details: Repeated eyewitness recall leads to reminiscence but not hypermnesia. *Journal of Applied Psychology, 79*, 260-271.
Wise, R. A., & Safer, M. A. (2004). What US judges know and believe about eyewitness testimony. *Applied Cognitive Psychology, 18*, 427-443.

Appendix

Questionnaire

Instructions for the refreshing phase;

It has been a while since you have seen the video with the accident and the story of the persons involved. Before we start asking more specific questions we like you to go back to the video. Please give a short and global description of the two story lines you have seen in the video. Details will be discussed later.

Cued recall questions;

One of the storylines in the video is about a male photo-model and his brother. The first scene of this storyline is about a photo shoot.

1A At the scene of the photo shoot the photo-model had a telephone conversation with his brother. Do you remember what they talked about?

1B Do you remember which persons, besides the photo-model himself, were present at the photo shoot?

The next scene shows the photo model and his brother at a car dealer.

2A Can you give a description of the clothes the brother of the photo-model was wearing in the scene at the auto dealer?

2B Can you give a description of the clothes the photo-model was wearing in the scene at the auto dealer?

The brother of the photo model invited his girlfriend for a diner at his house. The photo model arrives later on that evening.

3A Do you remember what the brother of the photo-model gets from his girlfriend when she arrived for diner?

3B Do you remember what the brother of the photo-model was wearing when his girlfriend arrived?

4A Do you remember what the conversation was about during dinner?

4B Do you remember who after dinner left the house first and why?

The photo model and his brother went back to the auto dealer for a test ride.

5A Do you remember the color of the car that the brother of the photo-model wanted to buy?

5B Do you remember the color of the car the brother and the photo-model took for a test ride?

The other storyline in the video shows an older couple. The first scene is at their house.

6A Do you remember what the older woman was doing while her husband was listening to music in the living room?

6B Do you remember what kind of music the older man was listening to?

The older man made reservations for a trip. His wife did not know about this.

7A Do you remember why the older woman began to suspect her husband of doing something behind her back?

7B Do you remember the destination of the trip the older man and woman planned to make?

8A Can you describe the first reaction of the older woman when she saw the suitcases here husband had packed?

8B Can you describe how and where the older man told his wife about the trip he booked?

The last part of the video shows the two brothers making a test ride en the older couple on their way to the airport by car.

9A Do you remember what problems the car had that the brother and the photo-model were driving for a test ride?

9B Do you remember the topic of conversation of the older couple during their ride to the airport?

During the test ride the car of the brothers passes two pedestrians narrowly.

10A Do you remember what the photo-model did immediately after passing the pedestrians?

10B Can you give a description of the pedestrians?

11A Was the older woman wearing a seatbelt during the ride to the airport?

11B Was the husband of the older female wearing a seatbelt during the ride to the airport?

12A Who was driving the car; the older woman or her husband?

12B Who was driving the car; the photo-model or his brother?

13A Which car drove through the red light?

13B Who is badly injured and carried away on a stretcher?

14A Can you describe how the two cars ran into each other?

14B Can you describe where and how the car with the photo-model came to a halt after the accident?

15A Can you describe the damage of the car of the older couple?

15B Can you describe the damage of the car of the photo-model?