

Author version of:

Krix, A. C., Sauerland, M., Gabbert, F., & Hope, L. (2014). Providing eyewitnesses with initial retrieval support: What works at immediate and subsequent recall? *Psychology, Crime & Law*, 20, 1005-1027.
doi:10.1080/1068316X.2014.902456

Providing eyewitnesses with initial retrieval support: What works at
immediate and subsequent recall?

Alana C. Krix, and Melanie Sauerland

Maastricht University

Fiona Gabbert

Goldsmiths University of London

Lorraine Hope

University of Portsmouth

Author Note

Alana C. Krix, Maastricht University, Faculty of Psychology and Neuroscience, P.O. Box 616, 6200 MD Maastricht, The Netherlands, Phone +31 43 3881613, E-mail Alana.Krix@maastrichtuniversity.nl; Melanie Sauerland, Maastricht University, Faculty of Psychology and Neuroscience, P.O. Box 616, 6200 MD Maastricht, The Netherlands, Phone +31 43 3881963. E-Mail Melanie.Sauerland@maastrichtuniversity.nl; Fiona Gabbert, Goldsmiths University of London, Department of Psychology, New Cross, SE14 6NW, United Kingdom, Phone

Running head: INITIAL RETRIEVAL SUPPORT FOR EYEWITNESSES

+44 20 70785496. E-Mail: f.gabbert@gold.ac.uk; Lorraine Hope, University of Portsmouth, Department of Psychology, King Henry I Street, Portsmouth, Hampshire PO1 2DY, United Kingdom, Phone +44 23 92846329, E-Mail lorraine.hope@port.ac.uk. Correspondence regarding this article should be addressed to Alana C. Krix.

Acknowledgments

We thank Jessica Arnet, Nils Hagemann, Annick Janssen, Stéphanie Kerkhofs, Brenda Nuñez-Gonzalez, Chiel Slegers, Tamara Sujew, and Brenda Urban for their help with data collection and Henry Otgaar for his valuable comments on an earlier version of the manuscript. We are also grateful to Timm M. Kuhl for preparing the data for analyses. This research was supported by a PhD scholarship from Studienstiftung des Deutschen Volkes (German National Academic Foundation) awarded to the first author.

Providing eyewitnesses with initial retrieval support: What works at immediate and subsequent recall?

The effect of retrieval support on eyewitness recall was investigated in two experiments. Based on the outshining hypothesis, Experiment 1 tested whether retrieval support enhances witness performance (compared to free recall) especially when witnessing conditions are suboptimal (e.g., because witnesses were distracted during the crime). Eighty-eight participants watched a videotaped crime with either full or divided attention and subsequently received retrieval support with the Self-Administered Interview[©] (SAI) or completed a free recall (FR). One week later (T2) all participants completed a second FR. Unexpectedly, retrieval support did not lead to better memory performance than FR under divided attention conditions suggesting that retrieval support is not effective to overcome adverse effects of divided attention. Moreover, presence of retrieval support at Time 1 (T1) had no effect on memory performance at T2. Experiment 2 ($N = 81$) tested the hypothesis that these T2-results were due to a reporting issue undermining the memory-preserving effect of T1-retrieval support by manipulating retrieval support (SAI vs. FR) at T1 and T2. As expected, T1-retrieval support led to increased accuracy at T2. Thus, the beneficial value of T1-retrieval support seems greatest with high-quality T2-interviews. Interviewers should consider this when planning a subsequent interview.

Keywords: eyewitness testimony, outshining hypothesis, self-administered interview, divided attention, retrieval support

Eyewitness testimony is an important element for successful police investigations, as it is often the only available evidence (Kebbell & Milne, 1998). Hence, obtaining the most complete and accurate statements from eyewitnesses is essential. In an effort to help witnesses remember, research on eyewitness testimony has identified various techniques that provide retrieval support. One such technique is the mental reinstatement of context which is part of the Cognitive Interview (CI, Fisher & Geiselman, 1992). The mental reinstatement of context mnemonic is based on the principle of encoding specificity, that is, the idea that retrieval should be facilitated if encoding context is recreated during retrieval (Tulving & Thomson, 1973). Specifically, witnesses are instructed to mentally recreate the context of encoding by thinking back to what they saw, heard, thought or felt during the incident. Such context cues are stored alongside the memory of the incident and facilitate retrieval by providing additional access pathways to memory.

Another example of an interview tool which provides ample retrieval support is the recently developed Self-Administered Interview[©] (SAI; Gabbert, Hope, & Fisher, 2009; see Hope, Gabbert, & Fisher, 2011, for a more detailed description of its structure and development process). It is a self-administered booklet designed to be completed by the witnesses directly after the crime to ensure an early interview when the police do not have the time and resources to conduct a timely personal interview. Thereby, it serves as a supplement, not a replacement, of a comprehensive subsequent personal interview. As a generic recall tool, it can be used for different kinds of crimes. The SAI is based upon the CI and adopts some of the CI's memory-enhancing components, such as the aforementioned mental context reinstatement and the report everything component. The latter comprises the instruction to provide the most complete and accurate account possible. Moreover, the SAI fosters a high-quality statement by discouraging witnesses from guessing and by using non-leading questions (Hope et al., 2011). The interview provides a structure for recall and consists of several sections each focusing on a different aspect of the testimony (e.g. course of events, or appearance of the perpetrator). Additionally, the retrieval of spatial information is supported, as witnesses are required to draw a sketch of the scene. To summarize, the SAI relies on multiple and varied (i.e., verbal and non-verbal) retrieval attempts which is

beneficial for recall because information that cannot be retrieved with one technique or retrieval attempt may well become accessible with another one (Tulving & Watkins, 1975).

Since the SAI is a relatively new tool, only few studies have been conducted to examine it so far, but those which have yielded promising results. Specifically, the SAI has been shown to elicit a more complete recall than a free recall (FR) without compromising accuracy (Gabbert et al., 2009, Experiment 1; Gawrylowicz, Memon, & Scoboria, 2013). FR is generally understood as a method that simply instructs people to provide a free narrative of what they can recall, without impeding, but also not supporting retrieval from memory. It lacks the memory-enhancing components of the SAI, such as the mental reinstatement of context and multiple and varied retrieval, and hence, provides only little retrieval support. More importantly, the SAI can preserve memory for a subsequent interview (Hope, Gabbert, Fisher, & Jamieson, in press). In Hope et al.'s study, participants were assigned to one of three groups. Participants initially completed either an SAI, an FR, or no interview at all (i.e., a no-initial-interview control group). In the second session one week later, all participants underwent a CI. Here the reports of the SAI group were more accurate than those of the other two groups. Interestingly, the FR group and the no-initial-interview control group did not differ in accuracy. This indicates that an early recall opportunity without additional retrieval support does not produce memory-enhancing effects at a subsequent retrieval attempt. Hope et al. (in press) explain these results with associative network models of memory (e.g., J. R. Anderson, 1983) according to which memory traces are represented as a network consisting of nodes and links between them. A high-quality retrieval attempt as implemented by the SAI with its ample retrieval support strengthens both the nodes and links of the network and thus facilitates retrieval in a subsequent recall attempt. The memory trace should be relatively less strengthened if the initial retrieval is less effortful, such as occurs with an FR.

It is the aim of the present study to examine the effect of retrieval support on eyewitness performance more closely in two experiments. Assuming a positive psychological perspective instead of focusing on deficits, the present paper seeks to increase knowledge regarding how to promote the reliability of eyewitness statements. Increasing this knowledge base further is important because of the critical role eyewitnesses often play during criminal

investigations and can lead to an adjustment of the existing interview practice. In Experiment 1, the potential positive effect of retrieval support on recall performance was examined when attention was divided during encoding. Experiment 2 took on the mixed results that were obtained in Experiment 1 and in previous research regarding the effect of providing retrieval support in the first interview on recall performance in subsequent interviews. Hence, we examined the conditions as to when providing witnesses with retrieval support at T1 can preserve memory at T2 and investigated whether the type of T2-interview matters in this regard. In both experiments, the SAI was selected as a proxy tool for providing witnesses with retrieval support. In doing so, we additionally contribute to increasing the available data base regarding the SAI. The FR was used as an interview format that does not provide retrieval support.

While observing a crime, witnesses may pursue a secondary task, such as speaking on the phone, paying a bill, driving a car, or looking for an escape route. As a consequence, their attention may not be fully allocated to the incident. Divided attention during encoding is suboptimal, as it has disruptive effects on encoding and can lead to reduced memory performance (e.g., N. D. Anderson, Craik, & Naveh-Benjamin, 1998; Craik, Govoni, Naveh-Benjamin, & Anderson, 1996). In the context of eyewitness testimony, divided attention has been found to increase witnesses' suggestibility (Lane, 2006) and is thought to play a role in the weapon focus effect which refers to a reduced memory performance if a weapon was present during the crime (Fawcett, Russell, Peace, & Christie, 2013). The negative effects of divided attention have clearly been documented. Yet, to our knowledge, the potential positive effect of providing witnesses with retrieval support on recall performance when they were distracted during the crime has not been investigated.

Observations in the domain of context-dependent memory suggest that retrieval support should be especially effective when suboptimal conditions, such as divided attention, were prevalent during the crime. The outshining hypothesis (Smith, 1988, 1994) predicts that environmental context cues are not always equally effective for memory retrieval. Specifically, if there are strong non-contextual memory cues during retrieval, the meaning of environmental context cues is relatively diminished (i.e., outshone). On the other hand, if

memory cues are weak, witnesses may profit relatively more from contextual cues when attempting to retrieve information. For example, Smith, Glenberg, and Bjork (1978; but see Smith & Vela, 2001) found that administering a recall test in the very room where the stimuli had been encoded, instead of in a different room, led to better memory performance. As predicted by the outshining hypothesis, no such difference was found for recognition performance: Here the to-be-recognized word was a strong enough non-contextual cue that outshone the context cues provided by the room. More support for the outshining hypothesis comes from Fernandez and Glenberg (1985). In an associative-processing task, their participants were presented with pairs of words and asked to generate sentences that included those word pairs. In a subsequent memory test, environmental context (same vs. different room) did not have an effect on memory performance. Hence, associative processing during encoding reduced the influence of context cues during recall (see also Smith & Vela, 2001).

Divided attention, relative to full attention, also diminishes associative processing during encoding (e.g., Castel & Craik, 2003; but see Naveh-Benjamin, Guez, & Marom, 2003). Because pursuing a secondary task consumes additional cognitive resources, building of meaningful associations between the information seen may be impeded. In line with the outshining hypothesis (Smith, 1988, 1994), the meaning of context cues for recall should be heightened in this suboptimal situation. Consequently, retrieval support (e.g., with the SAI) may enhance recall more than an FR in suboptimal conditions, because it provides mental context cues for recall.

With these considerations in mind, the aim of Experiment 1 was to test whether providing witnesses with retrieval support can limit the detrimental effects of divided attention conditions. Participants completed two recall attempts, one immediately after witnessing a staged crime (T1) and one a week later (T2). For T1, we expected the reports of retrieval support participants to be more complete than those of FR participants, without compromising accuracy. In line with the outshining hypothesis (Smith, 1988, 1994), we expected an interaction between the amount of attention paid and the presence of retrieval support. Specifically, the decline in memory performance from full to divided attention should be less marked in the retrieval support group, due to the beneficial effect of the mental

context reinstatement instruction. Therefore, the difference between the retrieval support and the FR group should be greater under divided than under full attention conditions. For T2, the same interaction between initial retrieval support and attention was expected as for T1. Apart from that and in line with network models of memory (e.g., J. R. Anderson, 1983), we expected to replicate the memory preserving-effect of the SAI, that is, of an interview providing ample retrieval support (Hope et al., in press). In other words, for T2, we expected that participants who were given retrieval support at T1 would provide both a more comprehensive and more accurate account than participants who completed an FR at T1.

Experiment 1

Method

Participants and Design

In total, 93 participants took part in Experiment 1. Five were excluded, because they did not attend the T2-session, leaving $N = 88$ participants for analysis (65 women; 41 German native speakers, 47 Dutch native speakers; age 18 to 64, $M = 21.9$, $Mdn = 20.5$ years). Sample size was based on the size in previous SAI research (Gabbert et al., 2009) and on a power analysis with G*Power 3.1 (Faul, Erdfelder, Lang, & Buchner, 2007). This analysis yielded a required total sample size of $N = 90$, given $\beta = .80$, $\alpha = .05$, and a medium to large effect size. Participants were mainly (95.5%) psychology undergraduates receiving course credit in exchange for participation or members of the general public. Performance did not differ between students and non-students. Participants were randomly assigned within a 2 (presence of retrieval support at T1: FR vs. SAI) x 2 (attention during encoding: divided vs. full) between-participants design and tested individually.

Materials

Film.

The non-violent stimulus film, presented without audio, lasted 3:14 min and depicted the staged theft of a laptop. Six amateur actors (4 men, 2 women; ages 21 - 36) appeared in the film. The incident took place in a university environment and showed a number of students studying in a communal area. One student left his laptop unattended while taking a phone call, and a thief, along with an accomplice, stole the laptop.

Divided Attention Task.

An arithmetic verification task adopted from Logie, Maylor, Della Sala, and Smith (2004) was selected as the divided attention task, as solving arithmetic problems has been shown to consume working-memory resources (Logie, Gilhooly, & Wynn, 1994). Twenty arithmetic problems and their corresponding solutions were presented via loudspeakers. The arithmetic problems, containing an equal number of addition, subtraction, multiplication, and division problems, were taken from a previous study (Jamieson, Gabbert, Allan, & Carson, 2009) and were presented at a rate of one problem every 10 s. One incorrect solution was presented for each of the four basic arithmetic operations. The participants' task was to report if they detected a solution to an arithmetic problem was incorrect. The onset of the presentation of the arithmetic problems was synchronized with the beginning of the film.

In this task, participants made $M = 3.20$ ($SD = 1.76$, $Mdn = 3$) errors, including both commission and omission errors. Accuracy of this task was not correlated with the interview performance measures at T1 or T2, $r_s(42) \leq .21$, $p_s \geq .179$.

Interviews.

Self-Administered Interview.

German and Dutch translations of the original English version of the SAI (Gabbert et al., 2009; see Hope et al., 2011, for a detailed description) were prepared using a back-translation process. In the SAI, before writing down their recollections, witnesses first

mentally reinstated the context. That is, they were instructed to think back to the witnessed incident and picture in their minds what they saw, what they were thinking and how they were feeling at the time. In separate sections, they were asked to provide descriptions of the course of events, the appearance of the perpetrator(s), and, if applicable, of potential other witnesses or vehicles involved. The perpetrator and vehicle sections featured non-leading recall prompts to cue recall. Witnesses were also requested to draw a sketch of the scene, so as to facilitate the retrieval of spatial information. Thereby, the SAI relied on multiple and varied retrieval. In a final section, the witnessing conditions were prompted to assess how well the witness had seen the incident (e.g., What time of day did the event occur?, Were there any obstructions in your view?). Throughout the whole interview, witnesses were instructed to provide the most complete and accurate account of the witnessed incident possible, but not to guess.

Free Recall.

Following Gabbert et al. (2009), the FR instructed participants to report all details that they could remember about the sequence of actions and events, and of all persons involved, including the perpetrator(s) and other witnesses. Analogous to the SAI, participants were instructed to provide the most accurate and complete account possible, but were discouraged from guessing. Note that the FR differed in significant ways from the SAI. First, it lacked significant memory-enhancing components (e.g., mental context reinstatement). Second, it entailed only one instead of multiple and varied retrieval attempts. Finally, unlike the SAI, the FR did not feature recall prompts to cue recall.

Procedure

The experiment consisted of two sessions. Using a cover story, we told participants that the study was about social perception. Upon arrival to the T1-session, participants signed an informed consent and provided their demographic data. In the full attention condition, participants were told that they would watch a film about a social situation. They were asked to watch the film carefully, because they would be required to answer some questions about it. No information was revealed about the nature of these questions. Participants in the

divided attention condition were additionally instructed to simultaneously listen to arithmetic problems carefully. Participants were to say “false” aloud if they considered a solution to be incorrect and to remain silent if they considered a solution to be correct. The film was then shown to participants on a 22 in. (55.88 cm) wide screen with a resolution of 1920 x 1080. The participants in the divided attention group simultaneously listened to the arithmetic problems and responses were recorded by the experimenter. After the film, participants completed a number of unrelated filler tasks for 30 min.

To reduce any effects of physical context on recall performance, participants were led to a different room prior to providing information on what they had seen in the film. Depending on the retrieval support condition, the participants were either handed the SAI or the FR. Participants provided the statements in their native language. No time limits were imposed on participants in any condition.

The T2-session was scheduled one week later. Participants were not told beforehand about the content of this session. As in Gabbert et al. (2009, Experiment 2), all participants received an FR form for providing their second statement. Instructions reminded them that they had become a witness to an incident in the previous week and that they would be asked to give a second statement about the incident. Hereafter, the participants were thanked for participation and dismissed. After the completion of data collection, all participants were fully debriefed.

Coding

Participants’ statements were transcribed and coded employing a comprehensive coding scheme consisting of 560 details, of which approximately 58% were person, 17% action, 11% object, and 13% setting details. A detail was entered into the coding system when the coders agreed on its inclusion. The statement “The thief (1) sat (2) at the rightmost (3) table (4).” would yield four details (see Wright & Holliday, 2007, for a similar approach). Subjective responses, such as “He was ugly.”, were not scored. A detail was considered correct, if it matched the content of the stimulus film, and considered incorrect, if it did not match the content of the film. Details were considered confabulated when they were both

incorrect and non-existent (e.g., the thief attacked the victim; see Dando, Wilcock, & Milne, 2009, for a similar approach). To code the accuracy of age, height, and weight estimates, we accepted deviations of plus or minus 2 years, 4 cm, or 3 kg, respectively, from the true value (see Fahsing, Ask, & Granhag, 2004, for a similar approach). The Dutch and German statements were coded by native speakers of each language.

Inter-Coder Reliability

To establish inter-coder reliability, the randomly selected statements of ten Dutch and ten German participants (i.e., 20 participants or 23% of the total sample) were independently coded by two coders. For correct recall, Cohen's κ were .99 and .98 for the German and Dutch statements, $ps < .001$, respectively. For incorrect recall, κ coefficients were .98 and .94, $ps < .001$, respectively.

Results and Discussion

An *alpha* level of .05 was used for all statistical tests. We report Cohen's *d* (Cohen, 1988) for dependent or independent samples for the main effects (with $df = 1$ in the numerator) and η_p^2 for the interaction effects (see Sporer & Cohn, 2011). To investigate recall performance as a function of presence of retrieval support at T1 (FR vs. retrieval support with the SAI) and attention (divided vs. full), we calculated two-way ANOVAs. The number of correct and incorrect details and accuracy (number of correct details divided by all reported details; see Meissner, Sporer, & Susa, 2008) were the dependent variables. Table 1 displays the means and standard deviations of all dependent variables for T1 and T2.¹

Insert Table 1 about here

Recall at T1

Number of correct details.

The main effects of retrieval support, $F(1, 84) = 15.25, p < .001, d = 0.60$, and attention, $F(1, 84) = 78.20, p < .001, d = 1.72$, were qualified by a significant interaction, $F(1, 84) = 4.44, p = .038, \eta_p^2 = 0.05$. Unexpectedly, the simple main effects analyses revealed that only within the full attention group did the retrieval support participants ($M = 113.41$) recall significantly more correct details than the FR participants ($M = 78.32$), $F(1, 84) = 18.07, p < .001, d = 1.18$. This was not the case when attention was divided (retrieval support: $M = 49.50$; FR: $M = 39.00$), $F(1, 84) = 1.62, p = .207, d = 0.42$.

Number of incorrect details.

At T1 the retrieval support group ($M = 13.73$) recalled more incorrect details than the FR group ($M = 7.55$), $F(1, 84) = 23.91, p < .001, d = -1.05$. The main effect of attention, $F(1, 84) < 0.01, p = .971, d = -0.01$, and the interaction, $F(1, 84) < 0.01, p = .971, \eta_p^2 < 0.01$, were non-significant.

Accuracy.

Accuracy of the T1-statements did not differ as a function of retrieval support, $F(1, 84) = 1.80, p = .184, d = -0.25$. However, full attention led to greater accuracy ($M = 89.87\%$) than divided attention ($M = 79.62\%$), $F(1, 84) = 29.41, p < .001, d = 1.16$. The interaction was non-significant, $F(1, 84) = 0.02, p = .895, \eta_p^2 < 0.01$.

Replicating previous findings (Gabbert et al., 2009; Hope et al., in press), at T1 the group receiving retrieval support with the SAI recalled more completely than the FR group (for the number of correct details, this only applied to the full attention condition). As previously found, this was not accompanied by a decrease of accuracy. However, the expected interaction pattern between attention and presence of retrieval support did not emerge. For the number of correct details, the retrieval support group only performed better than the FR group when participants had watched the film with full attention.

Recall at T2

Contrary to our expectations, retrieval support at T1 did not have any effects on recall performance at T2, $F_s(1, 84) \leq 0.02, p_s \geq .902, d_s = 0.02$. The interactions between retrieval support and attention were also non-significant, $F_s(1, 84) \leq 1.68, p_s \geq .198, \eta_p^2 \leq 0.02$. The

full attention group ($M = 75.80$) recalled more correct details at T2 than the divided attention group ($M = 34.86$), $F(1, 84) = 58.72$, $p < .001$, $d = 1.64$. Their reports were also more accurate ($M = 91.33\%$) than in the divided attention group ($M = 80.26\%$), $F(1, 84) = 27.60$, $p < .001$, $d = 1.13$. For incorrect recall, the main effect of attention was non-significant, $F(1, 84) = 0.20$, $p = .659$, $d = 0.10$. Hence, we did not obtain a memory-preserving effect of initial retrieval support which was found earlier (Hope et al., in press) and which would have been predicted by network models of memory (e.g., J. R. Anderson, 1983). To explore why we did not detect any differences at T2, we investigated memory performance over time.

Comparing T1 and T2 Recall

To analyze performance over time, we included Time (T1 vs. T2) as a within-participant factor, yielding a $2 \times 2 \times 2$ mixed-factorial analysis. For reasons of brevity, we will not report the main effects of attention. All other main effects and interactions not mentioned in following were non-significant, $F_s(1, 84) \leq 3.27$, $p_s \geq .074$, $|d|_s \leq 0.11$, η_p^2 s ≤ 0.04 . For testing why there were no effects of T1-retrieval support on performance at T2, the interaction effects between time and presence of retrieval support are most relevant.

Number of correct details over time.

For the correct details, there were significant main effects of time, $F(1, 84) = 98.64$, $p < .001$, $d = 0.38$, and retrieval support, $F(1, 84) = 4.73$, $p = .033$, $d = 0.34$, which were qualified by a significant interaction, $F(1, 84) = 55.71$, $p < .001$, $\eta_p^2 = 0.40$. While the number of correct details significantly dropped over time within the retrieval support group (T1: $M = 81.45$; T2: $M = 55.66$), $F(1, 86) = 129.81$, $p < .001$, $d = 0.59$, it remained stable within the FR group (T1: $M = 58.66$; T2: $M = 55.00$), $F(1, 86) = 2.61$, $p = .110$, $d = 0.12$. The results show that while the number of correct details was much higher in the retrieval support group than in the FR group at T1, over time the number of correct details of the retrieval support group fell to the level of the FR group.

Number of incorrect details over time.

For the incorrect details, there were also significant main effects of time, $F(1, 84) = 28.75$, $p < .001$, $d = 0.48$, and retrieval support, $F(1, 84) = 7.43$, $p = .008$, $d = -0.59$, qualified

by a significant interaction, $F(1, 84) = 31.91, p < .001, \eta_p^2 = 0.28$. While the number of incorrect details significantly dropped over time within the retrieval support group (T1: $M = 13.73$; T2: $M = 7.61$), $F(1, 86) = 61.81, p < .001, d = 0.95$, it remained stable within the FR group (T1: $M = 7.55$; T2: $M = 7.70$), $F(1, 86) = 0.04, p = .838, d = -0.03$.

Accuracy over time.

The interaction between Time and retrieval support was marginally significant, $F(1, 84) = 3.75, p = .056, \eta_p^2 = 0.04$. Accuracy of the retrieval support group significantly increased from T1 ($M = 83.47\%$) to T2 ($M = 85.91\%$), $F(1, 86) = 5.95, p = .017, d = -0.25$, but remained stable in the FR group (T1: $M = 86.01\%$; T2: $M = 85.69\%$), $F(1, 86) = 0.11, p = .746, d = 0.03$.

The aim of Experiment 1 was to examine the effect of retrieval support on recall performance when the witness' attention was divided during encoding. Based on the outshining hypothesis (Smith, 1988, 1994), we hypothesized that the retrieval support group would always show better recall performance than the FR group. More importantly, this advantage should be more marked when attention was divided than when it was full. Unexpectedly, this pattern of results did not emerge for either T1 or T2. Note, however, that when the interactions were non-significant, accepting the null hypothesis cannot be attributed to a lack of power, even though the observed statistical power was low (power $\leq 0.06, fs \leq 0.032, \alpha = .05$). Indeed, none of the means displayed the descriptive trend in the expected direction, as can be seen in Table 1. The same was true in the case where the interaction reached significance (power = 0.52, $fs = 0.228, \alpha = .004$). For the number of correct details at T1, the retrieval support group recalled more correct details than the FR group within the full attention condition. Descriptively, this difference was much smaller within the divided attention condition and recall performance did not differ as a function of the presence of retrieval support. Even though our hypothesis was not supported, it should be noted that the retrieval support group never performed worse than the FR group. It seems thus that providing witnesses with retrieval support is not an effective means to overcome the adverse effects of divided attention during encoding.

Regarding both the number of correct and incorrect details, the retrieval support group experienced a significant loss of reported details and fell precisely to the level of the FR group over time. The differences on a descriptive level were so small that at T2 no differences in recall performance as a function of the presence of T1-retrieval support occurred. Hence, accepting the null hypothesis cannot be attributed to a lack of power (power = 0.05, $f_s \leq 0.014$, $\alpha = .05$). It is striking that the retrieval support group was apparently unable to uphold their level of performance at T2. This contradicts the findings obtained by Hope et al. (in press) and shows that retrieval support at T1 does not always lead to carry-over effects at T2.

Interestingly, previous studies examining recall performance with the CI, which also provides ample retrieval support, report results similar to ours. In these studies, CI performance was compared with standard interview (SI; i.e., an interview with less retrieval support) performance on several recall attempts (e.g., Brock, Fisher, & Cutler, 1999; McCauley & Fisher, 1995; Memon, Wark, Bull, & Köhnken, 1997). Completing the CI at T1, relative to the SI, did not influence memory performance at T2 (i.e., there were no carry-over effects from T1 to T2), irrespective of the type of T2-interview used. Brock et al. (1999) orthogonally manipulated type of interview at T1 and T2 (CI vs. SI). Over time, the greatest number of details was forgotten when participants first received the CI and then the SI. Strikingly, we also found the greatest loss of reported information over time in the retrieval support/FR condition. Possible reasons given for the failure to find carry-over effects included that memory may become less context-dependent over time (Memon et al., 1997) or that the CI facilitates non-retrieval processes (e.g., the communication of recollections) instead of memory retrieval itself (Brock et al., 1999). Yet, none of the studies followed up on these assumptions, so the reasons as to why no memory-preserving effect occurred still remain unclear.

To explain the difference in results regarding the current and Hope et al.'s (in press) experiment, we focused on the main difference between the two studies, that is, the interviews used at T2. Hope et al. who found a memory-preserving effect after retrieval support at T1 used a CI at T2, an interview with ample retrieval support, while we employed

an FR, an interview with little retrieval support. The memory-preserving effect of initially providing retrieval support may only be detectable when the T2-interview possesses certain features. It should be noted, though, that the use of an FR at T2 in the present study seemed justified based on both previous research results and applied considerations. Indeed, Gabbert et al. (2009, Experiment 2) also used an FR at T2 and found that participants who received initial retrieval support by completing an SAI recalled more correct details and provided more accurate reports in a T2-FR than a no-initial-interview control group. Moreover, there are countries, in which the CI is not used at all (e.g., in the Netherlands) or used very infrequently (e.g., in Germany) by the police. Therefore, we considered examining the effectiveness of initial retrieval support on the performance in T2-interviews with varying degrees of retrieval support (i.e., FR instead of CI as the T2-interview) to be highly relevant.

The question that remains to be answered is, whether a memory issue (i.e., actual forgetting of information), or a reporting issue (i.e., failure to overtly report details from memory) is accountable for the results of Experiment 1. Previous findings and theory render the first explanation unlikely. Retrieving information from memory increases the probability that this information will later be recalled (e.g., Ebbesen & Rienick, 1998; Foos & Fisher, 1988; McDaniel & Fisher, 1991), a phenomenon known as the *testing effect* (Roediger & Karpicke, 2006). According to network models of memory (e.g., J. R. Anderson, 1983), an early recall strengthens the network which results in a stronger memory trace. Moreover, it has been found that especially a high-quality initial recall can aid later recall attempts (e.g., Foos & Fisher, 1988; Marsh, Tversky, & Hutson, 2005; Tversky & Marsh, 2000). In Experiment 1, the SAI with its ample retrieval support elicited more correct details than the FR at T1 without affecting accuracy. Therefore, it should have led to better recall performance at T2, regardless of the T2-interview type.

Instead of a memory issue, we believe it likely that a reporting issue is responsible for the results. In their metamemory model, Koriat and Goldsmith (1996) distinguish between recall of information and overt reporting of this information as two separate components. People do not report everything they can recall. Instead, reporting details from memory is considered an active process which is influenced by one's current personal and situational

goals as well as task demands (Bartlett, 1932; Neisser, 1996). In free report conditions, as in the present experiment, participants have the freedom to report or withhold information (Koriat & Goldsmith, 1994). Koriat and Goldsmith (1996) termed this the *control of report option*. In line with this, the retrieval support group may have withheld relatively more information than the FR group and, as a consequence, may have underreported at T2. The reason for this might be that the retrieval support group chose not to report certain details, because they “only” received FR instructions and did not assume that an elaborate account was expected of them. The sparse FR at T2 that provided little structure and retrieval support (relative to the SAI) may thus have led to a provision of only a bare minimum of details.

At this time, results regarding the effect of early retrieval support on subsequent recall performance are mixed (Brock et al., 1999; Hope et al., in press; Memon et al., 1997) and knowledge about the conditions regarding when early retrieval support preserves memory is limited. Yet, witnesses are likely to be interviewed on several occasions in the course of the investigations (Bornstein, Liebel, & Scarberry, 1998). Therefore, it is important to study which measures need to be taken to ensure that recall is as complete and accurate as possible in both the first and subsequent interviews. Hence, we followed the results of Experiment 1 up to shed light onto the conditions during which providing retrieval support at T1 can preserve memory for a subsequent interview. Here we examined the effect of initial retrieval support on subsequent recall by orthogonally manipulating the presence of retrieval support (i.e., FR vs. retrieval support with the SAI) both at T1 and at T2. The FR at T2 was supplemented with non-leading cued questions. This interview (i.e., FR and subsequent cued questions) is quite common in cases where a CI is not regularly administered. Both types of interviews used at T2 were meant to incite the expectation to provide an elaborate account. Moreover, analyses of the FR at T2 group with and without the responses to the cued questions enabled us to make direct comparisons with the findings of Experiment 1. This approach allowed us to disentangle, whether a memory issue or reporting issue was responsible for the results found in Experiment 1.

For T1-recall, we hypothesized that the retrieval support group would recall more details than the FR group, while there would be no influence on accuracy. For T2-recall, we

expected that the retrieval support group would recall more details and at a higher accuracy than the supplemented FR (FR_{sup}) group. Furthermore, we predicted carry-over effects from T1 to T2-recall (i.e., a memory-preserving effect of initial retrieval support on subsequent recall). Specifically, participants who received retrieval support at T1, relative to the FR, were expected to report more details and at a higher accuracy one week later. Moreover, due to its beneficial effect, participants who received retrieval support at both T1 and T2 should be most accurate at T2. Participants who first had an FR and then an FR_{sup} should recall at the lowest accuracy, with the other two interview groups (retrieval support-FR_{sup}, FR-retrieval support) falling in-between.

Experiment 2

Method

Participants and Design

Eighty-seven native German participants took part in Experiment 2. Six participants were excluded from analyses due to non-attendance at T2, or a failure to comply with instructions, leaving 81 participants for analysis (59 women; age 19 to 33, $M = 22.22$, $Mdn = 21$ years). The majority (80.2%) were psychology undergraduates. Remaining participants were recruited in the vicinity of the university. Performance did not differ between psychology students and the other participants. In exchange for participation participants either received course credit or a €15 gift voucher. Experiment 2 employed a 2 x 2 between-participants design. Presence of retrieval support at T1 (FR vs. retrieval support with the SAI) and presence of retrieval support at T2 (FR_{sup} vs. retrieval support SAI) were the independent variables.

Materials

The stimulus film, filler tasks, and interviews were identical to Experiment 1. The non-leading cued questions that supplemented the FR at T2 comprised eight questions

regarding the incident, 14 questions regarding the appearance of the persons involved in the crime and one question concerning the appearance of other persons pictured (see Appendix).

Procedure

In Experiment 2 all participants watched the film with full attention and, depending on the T1-retrieval support condition, completed either an SAI or an FR at T1. One week later (T2) participants either received an SAI or an FR_{sup}. In the FR_{sup} condition, after providing their free reports, participants answered cued questions regarding the incident, the perpetrator(s) and other persons involved. Participants were instructed to answer the cued questions as completely and accurately as possible, but were discouraged from guessing. They should skip a question or respond with “I don’t know” if they did not know the answer. We administered the questions in the FR_{sup} at T2 group only, because we assumed that, unlike an SAI, an ordinary FR would not create the expectation to provide an elaborate recall account. The T1-FR was not followed by cued questions, because we wanted to use the same T1-interviews as in Experiment 1 for the sake of comparability.

Coding

For the FR_{sup} at T2 group all details mentioned in the free report were used to calculate the dependent variables. This was done both when a detail was consistently reported in the free report and the cued questions and when there were contradictions. Details elicited by the cued questions were only added when the respective details had not been mentioned in the preceding free report. Following this approach, $M = 23.63\%$ ($SD = 7.24$) of the correct details and $M = 47.69\%$ ($SD = 19.24$) of the incorrect details reported at T2 came from the cued questions. Hence, unless stated otherwise, participants’ performance in the FR at T2 includes the responses to the cued questions.

Results and Discussion

Table 2 displays the means and standard deviations of all dependent variables for T1 and T2.²

Insert Table 2 about here

Recall at T1

To investigate performance at T1, we calculated one-way ANOVAs with presence of T1-retrieval support as the independent variable. Correct and incorrect details and accuracy were the dependent variables.

Analogous to Experiment 1, at T1 the retrieval support group ($M = 119.17$) recalled more correct details than the FR group ($M = 90.05$), $F(1, 79) = 29.09$, $p < .001$, $d = 1.20$. The number of incorrect details at T1 was also larger for the retrieval support group ($M = 14.78$) than for the FR group ($M = 11.45$), $F(1, 79) = 7.92$, $p = .006$, $d = -0.63$. Accuracy did not differ as a function of T1-retrieval support, $F(1, 79) = 0.03$, $p = .867$, $d = 0.04$. Thus, in line with our hypotheses and similar to Experiment 1 and previous findings (Gabbert et al., 2009; Gawrylowicz et al., 2013; Hope et al., in press), we found that at T1 initial retrieval support elicited a more comprehensive recall than an initial FR, without compromising accuracy. Next, we examined recall performance at T2.

Recall at T2

T2 recall performance was analyzed using two-way ANOVAs with presence of T1-retrieval support and presence of T2-retrieval support as the independent variables. All analyses regarding the number of correct details at T2 were non-significant, $F_s(1, 77) \leq 1.92$, $ps \geq .170$, $|d|s \leq 0.31$, $\eta_p^2 = .02$. Other main effects or interactions not mentioned in the following were non-significant, $F(1, 77)s \leq 3.84$, $ps \geq .054$, $|d|s \leq 0.38$, $\eta_p^2 \leq 0.02$. The same holds for non-significant post-hoc comparisons with alpha adjustments, $ps \geq .039$, $ds \leq 0.64$.

Number of incorrect details.

Participants who received retrieval support at T1 ($M = 17.83$) recalled significantly fewer incorrect details one week later than participants who completed an FR at T1 ($M = 22.80$), $F(1, 77) = 10.28$, $p = .002$, $d = 0.65$. The T2-retrieval support group ($M = 16.85$) also

recalled fewer incorrect details than the T2-FR_{sup} group ($M = 23.80$), $F(1, 77) = 20.47$, $p < .001$, $d = 0.96$.

Accuracy.

Participants who received retrieval support at T1 ($M = 85.98\%$) recalled more accurately at T2 than participants who completed an FR at T1 ($M = 83.75\%$), $F(1, 77) = 5.53$, $p = .021$, $d = 0.49$. Likewise, the T2-retrieval support group ($M = 86.80\%$) provided a more accurate account than the T2-FR_{sup} group ($M = 82.90\%$), $F(1, 77) = 17.49$, $p < .001$, $d = 0.92$.

We also examined the effect of providing participants with retrieval support both at T1 and at T2 relative to the other interview combinations on recall accuracy. Descriptively, accuracy was highest in the group receiving retrieval support at both T1 and T2 ($M = 88.13\%$, $SD = 4.85$), followed by the FR/retrieval support group ($M = 85.41\%$, $SD = 3.49$), the retrieval support/FR_{sup} group ($M = 83.72\%$, $SD = 4.60$), and the FR/FR_{sup} group ($M = 82.09\%$, $SD = 3.45$). Overall, the difference between the groups was significant, $F(3, 77) = 7.91$, $p < .001$, $\eta_p^2 = 0.24$. As expected, post-hoc comparisons (with an adjusted Bonferroni correction of $\alpha = .017$; Shaffer, 1986) revealed that at T2, the group receiving retrieval support at both T1 and T2 recalled significantly more accurately than both the retrieval support/FR_{sup}, $p = .001$, $d = 0.93$, and the FR/FR_{sup} group, $p < .001$, $d = 1.43$. The FR/retrieval support group recalled more accurately than the FR/FR_{sup} group, $p = .014$, $d = 0.96$.

Consistent with Hope et al. (in press), we found a memory-preserving effect of initially providing participants with retrieval support: Participants who initially received retrieval support achieved a higher accuracy at T2 than participants who initially completed an FR. Somewhat unexpected, this was rather driven by the number of incorrect details than by the number of correct details. The beneficial effect of retrieval support is further revealed in the finding that participants who had retrieval support at both recall attempts recalled more accurately than participants who first completed an FR and then an FR_{sup}. The type of T2-interview is also important which is indicated by the recall of fewer incorrect details and higher accuracy of the T2-retrieval support group than the T2-FR_{sup} group.

Comparing T1 and T2 Recall

To further explore the pattern of results found at T2, we investigated memory performance over time, adding Time (T1 vs. T2) as a within-participants factor to the ANOVAs. In the following, we will only describe the significant results. The results of the ANOVAs can be found in Table 3.

Insert Table 3 about here

Number of correct details over time.

For correct recall, there were significant Time by T1-retrieval support and Time by T2-retrieval support interactions. Irrespective of the presence of T2-retrieval support, the number of correct details declined over time within the T1-retrieval support group (T1: $M = 119.17$; T2: $M = 109.44$), $F(1, 79) = 21.25$, $p < .001$, $d = 0.38$. In contrast, the number of correct details increased over time within the T1-FR group (T1: $M = 90.05$; T2: $M = 116.65$), $F(1, 79) = 154.91$, $p < .001$, $d = -1.23$. Irrespective of the presence of T1-retrieval support, providing participants with retrieval support at T2 did not have an influence on the number of correct details over time, $F(1, 79) = 1.48$, $p = .227$, $d = -0.16$. In contrast, completing an FR_{sup} at T2 resulted in an increase in the number of correct details (T1: $M = 102.78$, $SD = 29.11$; T2: $M = 115.03$, $SD = 21.85$), $F(1, 79) = 11.89$, $p = .001$, $d = -0.46$.

Number of incorrect details over time.

For incorrect recall, there were also significant Time by T1-retrieval support and Time by T2-retrieval support interactions. Irrespective of the presence of T2-retrieval support, the number of incorrect details increased over time within the T1-retrieval support group (T1: $M = 14.78$; T2: $M = 17.83$), $F(1, 79) = 9.02$, $p = .004$, $d = -0.41$. Yet, within the T1-FR group the number of incorrect details doubled over time (T1: $M = 11.45$; T2: $M = 22.80$), $F(1, 79) = 121.99$, $p < .001$, $d = -1.80$.

Providing retrieval support at T2 resulted in an increase in the number of incorrect details over time, irrespective of the presence of T1-retrieval support (T1: $M = 13.05$, $SD = 4.93$; T2: $M = 16.85$, $SD = 6.95$), $F(1, 79) = 12.33$, $p = .001$, $d = -0.62$. Completing an FR_{sup}

at T2 led to an even greater increase in incorrect details over time (T1: $M = 13.23$, $SD = 6.19$; T2: $M = 23.80$, $SD = 7.53$, $F(1, 79) = 92.89$, $p < .001$, $d = -1.52$).

Accuracy over time.

For accuracy, there were also significant Time by T1-retrieval support and Time by T2-retrieval support interactions. Irrespective of the presence of T2-retrieval support, accuracy decreased over time within the T1-retrieval support group (T1: $M = 88.87\%$; T2: $M = 85.98\%$), $F(1, 79) = 23.10$, $p < .001$, $d = 0.60$. An even more marked decrease over time occurred within the T1-FR group (T1: $M = 88.72\%$; T2: $M = 83.75\%$), $F(1, 79) = 66.90$, $p < .001$, $d = 1.22$.

Similar results were obtained regarding the presence of T2-retrieval support. Irrespective of the presence of T1-retrieval support, accuracy decreased over time within the T2-retrieval support group (T1: $M = 88.95\%$, $SD = 3.87$; T2: $M = 86.80\%$, $SD = 4.41$), $F(1, 79) = 14.90$, $p < .001$, $d = 0.51$. The decrease of accuracy over time was even stronger within the T2-FR_{sup} group (T1: $M = 88.68\%$, $SD = 4.28$; T2: $M = 82.90\%$, $SD = 4.10$), $F(1, 79) = 104.76$, $p < .001$, $d = 1.38$.

In summary, although memory performance declined from T1 to T2, retrieval support at T1 relative to an FR seemed to limit memory deterioration: Providing participants with retrieval support at T1 led to a smaller increase in incorrect details and a smaller decrease of accuracy over time than an FR at T1. This is in line with a memory-preserving effect of providing initial retrieval support (Hope et al., in press).

Recall performance at T2 without the cued questions.

To further clarify the question, whether a reporting issue was responsible for the results of Experiment 1, we analyzed recall performance at T2 without the responses to the cued questions that supplemented the FR at T2. Taking away those responses yields an FR which precisely corresponds to the T2-interview used in Experiment 1. A replication of the null findings in these analyses would support the reporting issue explanation of Experiment 1. Non-replication of the findings would support the memory issue explanation.

Apart from a marginally significant main effect of presence of T1-retrieval support on the number of incorrect details, $F(1, 77) = 3.84, ps = .054, ds = 0.38$, presence of T1-retrieval support did not have significant effects on recall performance at T2, $F_s(1, 77) \leq 1.85, ps \geq .178, ds \leq 0.30$. Note that this differs from the analyses reported above that included the cued recall. The main effect of the presence of T2-retrieval support was significant for both correct, $F(1, 77) = 28.50, p < .001, d = 1.16$, and incorrect details, $F(1, 77) = 17.04, p < .001, d = -0.85$, but not accuracy, $F(1, 77) = 2.82, p > .097, d = -0.36$. The main effects were qualified by significant T1-retrieval support x T2-retrieval support interactions (correct details: $F(1, 77) = 4.76, p = .032, \eta_p^2 = 0.06$; incorrect details: $F(1, 77) = 9.22, p = .003, \eta_p^2 = 0.11$). Importantly, simple main effects analyses revealed that providing participants with retrieval support at T1 only had an influence on T2-recall performance when there was also retrieval support at T2. Compared to the group receiving retrieval support at both T1 and T2 ($M = 104.43$), the FR/retrieval support group ($M = 117.95$) tended to recall more correct details, $F(1, 77) = 3.73, p = .057, d = -0.57$. The FR/retrieval support group also recalled more incorrect details at T2 ($M = 20.30$) than the group receiving retrieval support at both T1 and T2 ($M = 13.57$), $F(1, 77) = 12.63, p = .001, d = 1.10$. Interestingly, as in Experiment 1, there was no effect of the presence of T1-retrieval support when the interview at T2 was an FR, $F_s(1, 77) \leq 1.34, ps \geq .251, |d|s \leq 0.39$. As can be seen in Table 2, after the removal of the responses to the cued questions, a similar decline of the number of reported details as in Experiment 1 was visible for the participants who first received retrieval support and subsequently completed the FR.

To sum up the results of Experiment 2, as expected, we found that initially providing witnesses with retrieval support preserves memory for a subsequent recall. That is, participants who received retrieval support relative to an FR at T1 recalled fewer incorrect details and hence provided more accurate accounts one week later. Moreover, the deterioration of memory over time was decelerated for the T1-retrieval support group relative to the T1-FR group. Note that these findings differ from the results obtained in Experiment 1. They are, however, consistent with the results obtained by Hope et al. (in press) and with network models of memory (e.g., J. R. Anderson, 1983), according to which an early recall

opportunity preserves memory. This especially holds for high-quality initial recall opportunities (e.g., Ebbesen & Rienick, 1998; Foos & Fisher, 1988; Tversky & Marsh, 2000) that support retrieval, such as the SAI.

Because we used the same types of interview at T1 across experiments, the discrepant findings are very likely to be related to the type of interviews used at T2. Such an interpretation is supported by additional data analyses of T2-recall in Experiment 2 that took into account the details mentioned in the FR only and excluded details elicited by the cued questions. By removing those details, the T2-FR of Experiment 2 was matched the T2-interview of Experiment 1. In these additional analyses, we replicated the results from Experiment 1. Specifically, there was no significant effect of the presence of T1-retrieval support on subsequent recall when an FR was applied at T2.

These results support our hypothesis that a reporting issue was responsible for the failure to detect any differences at T2 across retrieval support conditions in Experiment 1. We assume that the different interviews used at T2 in Experiment 1 vs. 2 produced different expectations about reporting. In free recall conditions, people can decide which information to withhold and which to report (Koriat & Goldsmith, 1994, 1996). This *control of report option* is influenced by current goals and task demands, such as the assumption that a complete account of what one has witnessed is expected. Accordingly, in Experiment 1 the retrieval support group may have withheld relatively more information than the FR group, because the scant free recall instructions did not generate the expectation that they should report a detailed and complete account at T2. Consequently, the retrieval support group may have underreported. This was not the case in Experiment 2 when participants received more elaborate recall instructions.

General Discussion

In two experiments, seeking to increase knowledge about ways to promote the reliability of eyewitness statements, we investigated the effect of initially providing witnesses with retrieval support on recall performance. For this purpose, we contrasted an interview

with a high amount of retrieval support, the recently developed SAI (Gabbert et al., 2009), with an interview that does not provide retrieval support, the FR. The aim of Experiment 1 was to test the effect of retrieval support when the witnessing conditions were suboptimal, because the witness was distracted during the crime. Previous research has comprehensively shown the negative effects of divided attention on memory performance (e.g., Craik et al., 1996; Lane, 2006), but little is known about ways in which to compensate its negative effects. Based on the outshining hypothesis (Smith, 1988, 1994), we expected retrieval support to be especially effective relative to an FR when the participants watched the stimulus film with divided vs. full attention. However, the predicted interaction pattern did not emerge. Note, however, that the retrieval support group did not perform worse than the FR group in any condition. These findings have relevance for interview practice in real cases. Specifically, the police should inquire whether the witness may have been distracted during the crime before administering the initial interview, because the advantage of providing retrieval support could be diminished. Nevertheless, it is not harmful to provide witnesses with retrieval support in the initial interview when they were distracted during the crime. As earlier findings indicate (Gabbert et al., 2009, Experiment 2), this is preferable to a situation where no initial interview is carried out at all.

In Experiment 1, we observed an unexpected tremendous loss of details over time within the retrieval support group rather than a memory-preserving effect of initial retrieval support on T2-recall performance. Since witnesses are likely to be interviewed several times (Bornstein et al., 1998), it is important to know when initial retrieval support can preserve memory for subsequent recall. Therefore, we addressed this question in Experiment 2 by orthogonally manipulating the presence of retrieval support at T1 and T2. Here we found a memory-preserving effect of initial retrieval support. The pattern of results supported our hypothesis that a reporting issue could explain the non-significant results of Experiment 1 (Koriat & Goldsmith, 1996). Apparently, the sparse FR at T2 used in Experiment 1 led to a provision of only a bare minimum of details, thus undermining the memory-preserving effect of initial retrieval support. From an applied perspective, the results clearly show that police investigators should be aware that witnesses about to testify need to know that a

comprehensive account is expected from them, otherwise an initial interview with ample retrieval support is unlikely to have the desired effect on the subsequent recall attempt. This could, for example, be established during the rapport building phase between interviewer and interviewee (Collins, Lincoln, & Frank, 2002; Vallano & Schreiber Compo, 2011). Our data also show that the type of interview at T2 is important for recall performance, even after initial retrieval support was provided, as the T2-retrieval support group recalled more accurately than the T2-FR_{sup} group. In summary, this means that an initial comprehensive retrieval support interview assists the police during the investigations immediately after a crime by preserving memory while this does not exempt police from applying best practice during the subsequent personal interview.

One remark on the difference in accuracy at T2 as a function of T1-retrieval support in Experiment 2 seems in order. Although inspection of the means suggests a rather small difference between the interview groups (about 2.2%) it should be noted that the respective effect size was moderate. Moreover, the difference in T2-accuracy as a function of presence of T1-retrieval support was driven by a large increase in incorrect details in the T1-FR group relative to the T1-retrieval support group. Specifically, the former group recalled almost 30% more incorrect details at T2 than the latter. Given that every single incorrect detail can entail dire consequences, this again highlights the importance of initial retrieval support.

Another issue worth mentioning concerns the large standard deviations we observed for both retrieval support and FR groups in both experiments, especially regarding the number of correct details. This indicates that there are substantial differences in the potential of the interviews to elicit information from participants. Given the homogeneity of each sample, this is unlikely to reflect individual differences in memory performance alone, because undergraduates are not likely to differ much in this cognitive task (see Peters, Jelicic, Verbeek, & Merckelbach, 2007). Rather, this could reflect differences in the control of the report option, that is, the readiness to report details from memory (Koriat & Goldsmith, 1996). That is, some participants may have a stricter response criterion than others. As a consequence, the former will provide fewer details than the latter. Future research should look into this issue.

Turning to the limitations of the present studies, one limitation pertains to the use of a stimulus film instead of a staged crime and the fact that it showed a rather non-arousing (i.e., not stress-inducing) crime. Another limitation applies to the use of an undergraduate sample rather than participants from the general public, with undergraduates most likely being more skilled at writing than the average witness. However, we do not think that any of these issues poses a threat to the validity of our results for the following reasons. In meta-analyses on the CI which is closely related to the SAI, medium of presentation (live vs. video) was either not found to be a moderator of recall performance (Memon, Meissner, & Fraser, 2010) or, contrary to what one would expect, effect sizes were larger when the event was staged than when a video was shown (Köhnken, Milne, Memon, & Bull, 1999). Therefore, we conclude that using a film probably did not influence our results or, if anything, led to an underestimation of the true effects. Meta-analytically, type of event (neutral vs. arousing) had a small, albeit significant, influence for correct recall such that arousing events yielded smaller effect sizes than neutral events (Memon et al., 2010). However, for both neutral and arousing events, the effect sizes in favor of the CI were very large ($d > 1$), suggesting that the use of a neutral event was unproblematic in the present study. Moreover, not all types of crime are likely to induce stress in the witness. Also, especially when there are multiple witnesses, there are probably bystanders who are not directly involved or watching the crime from a distance and hence, may not be stressed by the crime. Note that crimes with multiple witnesses are also the cases in which the SAI is most likely to be used (Hope et al., 2011). Still, we encourage future projects examining whether more arousing events interact with retrieval support with the SAI which would also increase and broaden the existing SAI database. Finally, results of a different SAI study suggest that the composition of the sample may not be decisive for the effectiveness of the SAI, at least for T1-recall. Using a sample recruited from the general public, Gawryłowicz et al. (2013) found that at T1 the SAI group recalled more correct details than the FR group, whereas accuracy did not differ. Note that in the present studies we found the same results for T1-recall performance and obtained similar effect sizes with an undergraduate sample. Future studies should nevertheless examine recall performance with the SAI across different witness populations.

Taken together, the results of the current experiments replicate and extend earlier findings concerning the effect of early retrieval support and are highly relevant for police interview practice. Specifically, like Gabbert et al. (2009) and Hope et al. (in press), we detected that providing retrieval support during the initial interview elicits a more comprehensive, albeit not more accurate, initial account than an FR. Unexpectedly, an interview with retrieval support is not more effective than an FR when the witness was distracted during the incident. Finally, reluctance to report was identified as an important determinant regarding whether a memory-preserving effect after conducting an initial interview with retrieval support emerges. The impact of an initial interview with ample retrieval support on a subsequent interview is greatest when the latter is a high-quality interview and when the witness assumes that a comprehensive account is expected.

References

- Anderson, J. R. (1983). A spreading activation theory of memory. *Journal of Verbal Learning and Verbal Behavior*, 22, 261-295. doi:10.1016/S0022-5371(83)90201-3
- Anderson, N. D., Craik, F. I. M., & Naveh-Benjamin, M. (1998). The attentional demands of encoding and retrieval in younger and older adults: 1. Evidence from divided attention costs. *Psychology and Aging*, 13, 405-423. doi:10.1037/0882-7974.13.3.405
- Bartlett, F. C. (1932). *Remembering: A study in experimental and social psychology*. Cambridge: Cambridge University Press.
- Bornstein, B. H., Liebel, L. M., & Scarberry, N. C. (1998). Repeated testing in eyewitness memory: A means to improve recall of a negative emotional event. *Applied Cognitive Psychology*, 12, 119-131. doi:10.1002/(SICI)1099-0720(199804)12:2<119::AID-ACP500>3.0.CO;2-4
- Brock, P., Fisher, R. P., & Cutler, B. L. (1999). Examining the cognitive interview in a double-test paradigm. *Psychology, Crime & Law*, 5, 29-45. doi:10.1080/10683169908414992
- Castel, A. D., & Craik, F. I. M. (2003). The effects of aging and divided attention on memory for item and associative information. *Psychology and Aging*, 18, 873-885. doi:10.1037/0882-7974.18.4.873
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Collins, R., Lincoln, R., & Frank, M. G. (2002). The effect of rapport in forensic interviewing. *Psychiatry, Psychology and Law*, 9, 69-78. doi:10.1375/pplt.2002.9.1.69
- Craik, F. I. M., Govoni, R., Naveh-Benjamin, M., & Anderson, N. D. (1996). The effects of divided attention on encoding and retrieval processes in human memory. *Journal of Experimental Psychology: General*, 125, 159-180. doi:10.1037/0096-3445.125.2.159

- Dando, C., Wilcock, R., & Milne, R. (2009). The Cognitive Interview: The efficacy of a modified mental reinstatement of context procedure for frontline police investigators. *Applied Cognitive Psychology, 23*, 138-147. doi:10.1002/acp.1451
- Ebbesen, E. B., & Rienick, C. B. (1998). Retention interval and eyewitness memory for events and personal identifying attributes. *Journal of Applied Psychology, 83*, 745-762. doi:10.1037/0021-9010.83.5.745
- Fahsing, I. A., Ask, K., & Granhag, P. A. (2004). The man behind the mask: Accuracy and predictors of eyewitness offender descriptions. *Journal of Applied Psychology, 89*, 722-729. doi:10.1037/0021-9010.89.4.722
- Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*, 175-191. doi:10.3758/bf03193146
- Fawcett, J. M., Russell, E. J., Peace, K. A., & Christie, J. (2013). Of guns and geese: A meta-analytic review of the 'weapon focus' literature. *Psychology Crime & Law, 19*, 35-66. doi:10.1080/1068316x.2011.599325
- Fernandez, A., & Glenberg, A. M. (1985). Changing environmental context does not reliably affect memory. *Memory & Cognition, 13*, 333-345. doi:10.3758/BF03202501
- Fisher, R. P., & Geiselman, R. E. (1992). *Memory enhancing techniques for investigative interviewing: The cognitive interview*. Springfield, IL: Charles C. Thomas.
- Foos, P. W., & Fisher, R. P. (1988). Using tests as learning opportunities. *Journal of Educational Psychology, 80*, 179-183. doi:10.1037//0022-0663.80.2.179
- Gabbert, F., Hope, L., & Fisher, R. P. (2009). Protecting eyewitness evidence: Examining the efficacy of a self-administered interview tool. *Law and Human Behavior, 33*, 298-307. doi:10.1007/s10979-008-9146-8
- Gawrylowicz, J., Memon, A., & Scoboria, A. (2013). Equipping witnesses with transferable skills: The Self-Administered Interview©. *Psychology, Crime & Law*. Advance online publication. doi:10.1080/1068316x.2013.777961

- Hope, L., Gabbert, F., & Fisher, R. P. (2011). From laboratory to the street: Capturing witness memory using the self-administered interview. *Legal and Criminological Psychology, 16*, 211-226. doi:10.1111/j.2044-8333.2011.02015.x
- Hope, L., Gabbert, F., Fisher, R. P., & Jamieson, K. (in press). Protecting and enhancing eyewitness memory: The impact of an initial recall attempt on performance in an investigative interview. *Applied Cognitive Psychology*.
- Jamieson, K., Gabbert, F., Allan, K., & Carson, D. (2009, July). *The effects of actual versus perceived memory quality on our susceptibility to memory conformity*. Paper presented at the meeting of the Society for Applied Research in Memory and Cognition, Kyoto, Japan.
- Kebbell, M. R., & Milne, R. (1998). Police officers' perceptions of eyewitness performance in forensic investigations. *The Journal of Social Psychology, 138*, 323-330. doi:10.1080/00224549809600384
- Köhnken, G., Milne, R., Memon, A., & Bull, R. (1999). The cognitive interview: A meta-analysis. *Psychology, Crime & Law, 5*, 3-27. doi:10.1080/10683169908414991
- Koriat, A., & Goldsmith, M. (1994). Memory in naturalistic and laboratory contexts: Distinguishing the accuracy-oriented and quantity-oriented approaches to memory assessment. *Journal of Experimental Psychology: General, 123*, 297-315. doi:10.1037//0096-3445.123.3.297
- Koriat, A., & Goldsmith, M. (1996). Monitoring and control processes in the strategic regulation of memory accuracy. *Psychological Review, 103*, 490-517. doi:10.1037/0033-295x.103.3.490
- Lane, S. M. (2006). Dividing attention during a witnessed event increases eyewitness suggestibility. *Applied Cognitive Psychology, 20*, 199-212. doi:10.1002/acp.1177
- Logie, R. H., Gilhooly, K. J., & Wynn, V. (1994). Counting on working-memory in arithmetic problem-solving. *Memory & Cognition, 22*, 395-410. doi:10.3758/bf03200866
- Logie, R. H., Maylor, E. A., Della Sala, S., & Smith, G. (2004). Working memory in event- and time-based prospective memory tasks: Effects of secondary demand and age.

European Journal of Cognitive Psychology, 16, 441-456.

doi:10.1080/09541440340000114

Marsh, E. J., Tversky, B., & Hutson, M. (2005). How eyewitnesses talk about events:

Implications for memory. *Applied Cognitive Psychology, 19*, 531-544.

doi:10.1002/acp.1095

McCauley, M. R., & Fisher, R. P. (1995). Facilitating children's eyewitness recall with the revised cognitive interview. *Journal of Applied Psychology, 80*, 510-516.

doi:10.1037//0021-9010.80.4.510

McDaniel, M. A., & Fisher, R. P. (1991). Tests and test feedback as learning sources.

Contemporary Educational Psychology, 16, 192-201. doi:10.1016/0361-

476X(91)90037-L

Meissner, C. A., Sporer, S. L., & Susa, K. J. (2008). A theoretical review and meta-analysis of the description-identification relationship in memory for faces. *European Journal of Cognitive Psychology, 20*, 414-455. doi:10.1080/09541440701728581

doi:10.1080/09541440701728581

Memon, A., Meissner, C. A., & Fraser, J. (2010). The cognitive interview: A meta-analytic review and study space analysis of the past 25 years. *Psychology, Public Policy, and Law, 16*, 340-372. doi:10.1037/a0020518

doi:10.1037/a0020518

Memon, A., Wark, L., Bull, R., & Köhnken, G. (1997). Isolating the effects of the cognitive interview techniques. *British Journal of Psychology, 88*, 179-197. doi:10.1111/j.2044-

8295.1997.tb02629.x

Naveh-Benjamin, M., Guez, J., & Marom, M. (2003). The effects of divided attention at encoding on item and associative memory. *Memory & Cognition, 31*, 1021-1035.

doi:10.3758/bf03196123

Neisser, U. (1996). Remembering as doing. *Behavioral and Brain Sciences, 19*, 203.

doi:10.1017/S0140525X00042308

Peters, M. J. V., Jelicic, M., Verbeek, H., & Merckelbach, H. (2007). Poor working memory predicts false memories. *European Journal of Cognitive Psychology, 19*, 213-232.

doi:10.1080/09541440600760396.

- Roediger, H. L., & Karpicke, J. D. (2006). Test-enhanced learning: Taking memory tests improves long-term retention. *Psychological Science, 17*, 249-255.
doi:10.1111/j.1467-9280.2006.01693.x
- Shaffer, J. P. (1986). Modified sequentially rejective multiple test procedures. *Journal of the American Statistical Association, 81*, 826-831. doi:10.1080/01621459.1986.10478341
- Smith, S. M. (1988). Environmental context dependent memory. In G. M. Davies, & D. M. Thomson (Eds.), *Memory in context: Context in memory* (pp. 13-33). New York: Wiley.
- Smith, S. M. (1994). Theoretical principles of context-dependent memory. In M. Gruneberg, & P. E. Morris (Eds.), *Theoretical aspects of memory* (pp. 168-195). London: Routledge.
- Smith, S. M., Glenberg, A. G., & Bjork, R. A. (1978). Environmental context and human memory. *Memory & Cognition, 6*, 342-353. doi:10.3758/BF03197465
- Smith, S. M., & Vela, E. (2001). Environmental context-dependent memory: A review and meta-analysis. *Psychonomic Bulletin & Review, 8*, 203-220. doi:10.3758/bf03197114
- Sporer, S. L., & Cohn, L. D. (2011). Meta-analysis. In B. D. Rosenfeld, & S. D. Penrod (Eds.), *Research methods in forensic psychology* (pp. 43-62). New York: Wiley.
- Tulving, E., & Thomson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review, 80*, 352-373. doi:10.1037/h0020071
- Tulving, E., & Watkins, M. J. (1975). Structure of memory traces. *Psychological Review, 82*, 261-275. doi:10.1037/h0076782
- Tversky, B., & Marsh, E. J. (2000). Biased retellings of events yield biased memories. *Cognitive Psychology, 40*, 1-38. doi:10.1006/cogp.1999.0720
- Vallano, J. P., & Schreiber Compo, N. (2011). A comfortable witness is a good witness: Rapport-building and susceptibility to misinformation in an investigative mock-crime interview. *Applied Cognitive Psychology, 25*, 960-970. doi:10.1002/acp.1789
- Wright, A. M., & Holliday, R. E. (2007). Enhancing the recall of young, young-old and old-old adults with cognitive interviews. *Applied Cognitive Psychology, 21*, 19-43.
doi:10.1002/acp.1260

Footnotes

¹The distributions of the number of incorrect details and accuracy of T1 and the number of correct and incorrect details, and accuracy of T2 were significantly skewed. However, the results of the analyses with the transformed variables yielded the same results. Therefore, we will report the analyses with the untransformed variables.

²The distribution of the number of incorrect details at T2 was significantly skewed. However, the results of the analyses with the transformed variable yielded the same results. Therefore, we will report the analyses with the untransformed variable.

Appendix: Cued Questions Following the Free Recall at T2 (Experiment 2)

Questions relating to the sequence of events.

1. How many people in total could be seen in the film?
2. Describe the crime scene!
3. What was stolen?
4. What precisely did the victim do so that the perpetrator/s could commit the theft?
5. Did the perpetrator/s have (an) accomplice/s? If yes, what did the accomplice/s do?
6. Describe how the perpetrator/s executed the theft!
7. Where did the perpetrator/s put the stolen object(s)?
8. If other people were present, what did they do during the theft?
9. Are there any further details regarding the course of events which you remember but which we have not asked you about? If yes, which details are these?

Questions relating to the appearance of the people involved in the crime (i.e., perpetrator and accomplice).

1. What was the sex of the person?
2. How old was the person?
3. How tall was the person in cm?
4. How heavy was the person in kg?
5. What was the ethnicity of the person?
6. Describe the figure of the person!
7. Describe the clothing of the person!
8. Did the person wear any jewelry or accessories? If yes, please describe!
9. What was the hair color of the person?

10. Describe the hair style and the hair length of the person!
11. Did the person have any distinctive features (e.g., scars, moles, or tattoos)? If yes, how did they look like?
12. Did the person wear glasses? If yes, how did they look like?
13. Did the person have facial hair? If yes, how did it look like?
14. Are there any other details about the person which you remember but which have not been asked about? If yes, which are these?

Question relating to all other people present in the film.

1. Were there any other people present in the film? If yes, please describe them.

Table 1

Means and Standard Deviations for Number of Correct and Incorrect Details, and Accuracy at Time 1 and Time 2 (Experiment 1)

	Time 1						Time 2					
	Divided attention		Full attention		Across attention conditions		Divided attention		Full attention		Across attention conditions	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Number of correct details												
FR	39.00	24.49	78.32	26.64	58.66	32.17	38.00	19.46	72.00	30.02	55.00	30.35
Retrieval Support	49.50	25.30	113.41	32.38	81.45	43.24	31.73	16.73	79.59	30.84	55.66	34.46
Across interviews	44.25	25.17	95.86	34.26	70.06	39.58	34.86	18.21	75.80	30.32	55.33	32.28
Number of incorrect details	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
FR	7.55	4.90	7.55	4.48	7.55	4.64	8.14	4.99	7.27	6.32	7.70	5.64
Retrieval Support	13.68	7.77	13.77	6.02	13.73	6.87	7.72	5.47	7.50	6.22	7.61	5.79
Across interviews	10.61	7.13	10.66	6.11	10.64	6.60	7.93	5.18	7.39	6.20	7.66	5.69
Accuracy	<i>M (%)</i>	<i>SD</i>	<i>M (%)</i>	<i>SD</i>	<i>M (%)</i>	<i>SD</i>	<i>M (%)</i>	<i>SD</i>	<i>M (%)</i>	<i>SD</i>	<i>M (%)</i>	<i>SD</i>
FR	81.01	14.39	91.01	4.75	86.01	11.74	80.55	14.29	90.82	6.51	85.69	12.14
Retrieval Support	78.22	7.48	88.73	5.39	83.47	8.35	79.97	10.67	91.84	5.52	85.91	10.32
Across interviews	79.62	11.42	89.87	5.15	84.74	10.21	80.26	12.46	91.33	5.99	85.80	11.20

INITIAL RETRIEVAL SUPPORT FOR EYEWITNESSES

Table 2

Means and Standard Deviations for Number of Correct Details and Incorrect Details, and Accuracy at Time 1 and at Time 2 (Experiment 2)

	Time 1		FR Time 2		FR _{sup} Time 2		Retrieval Support Time 2		Across interviews Time 2 ^a	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Number of correct details										
FR Time 1	90.05	22.55	80.50	14.96	115.35	17.57	117.95	22.66	116.65	20.06
Retrieval Support Time 1	119.17	25.89	88.70	25.46	114.70	25.91	104.43	24.83	109.44	25.58
Across interviews Time 1	104.79	28.24	84.60	21.02	115.03	21.85	111.02	24.48	113.00	23.16
Number of incorrect details										
FR Time 1	11.45	5.03	10.65	4.90	25.30	6.62	20.30	6.91	22.80	7.14
Retrieval Support Time 1	14.78	5.60	12.10	6.90	22.30	8.23	13.57	5.29	17.83	8.11
Across interviews Time 1	13.14	5.55	11.38	5.95	23.80	7.53	16.85	6.95	20.28	8.00
Accuracy	<i>M (%)</i>	<i>SD</i>	<i>M (%)</i>	<i>SD</i>	<i>M (%)</i>	<i>SD</i>	<i>M (%)</i>	<i>SD</i>	<i>M (%)</i>	<i>SD</i>
FR Time 1	88.72	4.27	88.45	4.60	82.09	3.45	85.41	3.49	83.75	3.81
Retrieval Support Time 1	88.87	3.78	88.48	5.07	83.72	4.60	88.13	4.85	85.98	5.18
Across interviews Time 1	88.80	4.01	88.46	4.78	82.90	4.10	86.80	4.41	84.88	4.66

Note. FR = free recall data including details from free report, but not cued questions; FR_{sup} = free recall data including both details from free report and cued questions.

^aThe data in the column “Across interviews Time 2” is the mean of FR_{sup} at Time 2 and retrieval support at Time 2.

INITIAL RETRIEVAL SUPPORT FOR EYEWITNESSES

Table 3

Results of the ANOVAs of the Analysis over Time (Experiment 2)

		<i>F</i>	<i>p</i>	Effect size	
				η_p^2	<i>d</i>
Number of correct details					
Main effects	Time	33.78	< .001	-	-0.31
	RS-T1	4.76	.032	-	0.48
	RS-T2	< 0.01	.989	-	< -0.01
Interactions	Time x RS-T1	154.10	< .001	0.67	-
	Time x RS-T2	6.66	.012	0.08	-
	RS-T1 x RS-T2	1.51	.223	0.02	-
	Time x RS-T1 x RS-T2	0.02	.882	< 0.01	-
Number of incorrect details					
Main effects	Time	134.23	< .001	-	-1.02
	RS-T1	0.39	.534	-	0.14
	RS-T2	8.61	.004	-	0.65
Interactions	Time x RS-T1	43.34	< .001	0.36	-
	Time x RS-T2	28.55	< .001	0.27	-
	RS-T1 x RS-T2	2.98	.088	0.04	-
	Time x RS-T1 x RS-T2	0.12	.735	< 0.01	-
Accuracy					
Main effects	Time	107.75	< .001	-	0.90
	RS-T1	1.88	.174	-	0.31
	RS-T2	6.14	.015	-	0.56
Interactions	Time x RS-T1	7.27	.009	0.09	-
	Time x RS-T2	22.20	< .001	0.22	-
	RS-T1 x RS-T2	1.06	.307	0.01	-
	Time x RS-T1 x RS-T2	0.63	.428	0.01	-

Note. $df_{num} = 1$; $df_{denom} = 77$. RS-T1 = factor presence of retrieval support at T1; RS-T2 = factor presence of retrieval support at T2.