Retrieval Enhances Eyewitness Suggestibility to Misinformation in Free and Cued Recall

Miko M. Wilford, M.S.

Jason C. K. Chan, Ph. D.

Sam J. Tuhn, B.S.

Iowa State University

Author Note

Miko M. Wilford, Department of Psychology, Iowa State University

Correspondence concerning this article should be sent to Miko M. Wilford, Department of Psychology, Iowa State University, W112 Lagomarcino, Ames, IA 50011-3180, USA

E-mail: mwilford@iastate.edu

Phone: (515) 294-9029

This material is based upon work supported by the National Science Foundation Graduate Research Fellowship under Grant. No. 202-18-94-00.

This is a manuscript of an article from *Journal of Experimental Psychology: Applied* 20 (2014): 81, doi:10.1037/xap0000001. Posted with permission.

Abstract (200 words)

Immediately recalling a witnessed event can increase people's susceptibility to later post-event misinformation. But this *retrieval-enhanced suggestibility* (RES) effect has only been shown when the initial recall test included specific questions that reappeared on the final test. Moreover, it is unclear whether this phenomenon is affected by the centrality of event details. These limitations make it difficult to generalize RES to criminal investigations, which often begin with free recall prior to more specific queries from legal officials and attorneys. In three experiments, we examined the influence of test formats (free recall versus cued recall) and centrality of event details (central versus peripheral) on RES. In Experiment 1, both the initial and final tests were cued recall. In Experiment 2, the initial test was free recall and the final test was cued recall. In Experiment 3, both the initial and final tests were free recall. Initial testing increased misinformation reporting on the final test for peripheral details in all experiments, but the effect was significant for central details only after aggregating the data from all three experiments. These results show that initial free recall can produce RES, and more broadly, that free recall can potentiate subsequent learning of complex prose materials.

Keywords: misinformation effect, eyewitness memory, investigative interviewing, false memory, testing effect

Retrieval Enhances Eyewitness Suggestibility to Misinformation in Free and Cued Recall

Misinformation is detrimental to the accuracy of eyewitness memory reports (Loftus & Palmer, 1974; Loftus, 1975; for a review, see Chrobak & Zaragoza, 2013). Because eyewitness testimony serves a crucial role in the legal system, discovering techniques that could reduce eyewitness suggestibility to misinformation is highly desirable. One potential technique is retrieval practice. The testing effect, which refers to enhanced retention due to retrieval practice (for a recent review, see Roediger, Putman, & Smith, 2011), has generated considerable research interest. If retrieval practice enhances memory of a witnessed event, then this enhancement would seem likely to reduce misinformation suggestibility. However, contrary to this possibility, retrieval practice can sometimes increase the negative influence of misleading suggestions on eyewitness memory performance — a finding termed retrieval-enhanced suggestibility (RES, Chan, Thomas, & Bulevich, 2009). Due to the frequency that witnesses engage in repeated retrieval, it is prudent to examine the conditions under which recalling details of a witnessed event might, or might not, produce RES.

Although the RES effect has been reported in various procedures (e.g., Chan & Langley, 2011; Chan & LaPaglia, 2011; 2013), several important aspects of its generality remain unclear. Specifically, all studies demonstrating RES possess two characteristics. First, the initial test was administered in a cued recall format, with questions querying specific details of the witnessed event. Second, event details contradicted by misinformation were mostly peripheral in nature (e.g., the type of car driven by a person, the color of the shirt worn by a person). The present research investigates whether initial retrieval increases suggestibility only under these conditions by examining a) whether RES occurs when the initial test is free recall and b) whether the RES effect occurs for both central and peripheral event details.

Applied Implications of Retrieval-Enhanced Suggestibility

From an applied perspective, it is crucial to understand whether the RES effect generalizes to situations in which initial recall does not resemble a cued recall test. Instead of asking questions about specific details of the witnessed event, criminal investigations often begin when witnesses call 9-1-1 and are asked to "tell everything," which is far more similar to a free recall test than a cued recall test. Indeed, the free recall report provided by an eyewitness during the early stages of an investigation can dictate how an investigation proceeds (e.g., if a witness recalls that the perpetrator is a white male, this information would constrain the subsequent search to suspects with similar characteristics), including what follow-up questions are asked by the investigators during subsequent interviews (e.g., if a witness recalls that the perpetrator got away in a car, the investigator might ask what type of car, etc.). Beyond reports to 911 operators, free recall is also prominently featured in all of the major standardized investigative interview protocols, including the Stepwise Interview (Yuille, Hunter, Joffe, & Zaparniuk, 1993), the Cognitive Interview (Fisher & Geiselman, 1992), the Achieving Best Evidence in Criminal Proceedings Interview (formerly Memorandum of Good Practice, Home Office, 1992, 2001), and the National Institute of Child Health and Human Development (NICHD) protocol (Lamb, 2007). Consequently, free recall plays a major role in criminal investigations. Because RES has only been demonstrated with an initial cued recall test, little is known about whether free recall retrieval practice would increase subsequent eyewitness suggestibility or produce a testing benefit. Indeed, if the RES effect is limited to situations for which initial recall involves the presentation of specific questions, then this constraint must be taken into account when considering the applicability of RES to criminal investigations. On the other hand, observing

RES following initial free recall would greatly increase its generality and potential applicability to criminal investigations.

Of additional importance to the generality and applicability of RES is the effect of item centrality. Memory for central details is typically less susceptible to misinformation than memory for peripheral details (Heath & Erickson, 1998; Paz-Alonso & Goodman, 2008, but see also Dalton & Daneman, 2006, who reported the opposite). In fact, some studies have demonstrated that memory for central details is not at all susceptible to the influence of misinformation (Loftus, 1979; Porter, Spencer, & Birt, 2003; Wright & Stroud, 1998). Currently, there is no general consensus on how best to define item centrality. Some researchers have defined central and peripheral information as details that are important and incidental to the thrust of the witnessed event, respectively (Dalton & Daneman, 2006; Loftus 1979; Wright & Stroud, 1998). Others defined centrality based on the frequency that those details are reported in free recall, with more frequently recalled items classified as central and less frequently recalled items as peripheral (Heath & Erickson, 1998; Saunders, 2009). We opted to define item centrality in the latter way because this definition is quantifiable and arguably more objective. If centrality had been defined by subjective norms applied by the experimenters or participants, it would be difficult to identify the precise criteria used to separate the central from the peripheral details. Further, subjective opinion regarding what is central versus peripheral to an event can vary greatly from one person to the next; this variation might have contributed to the disputed vulnerability of central details to misinformation—even among studies that adopted similar methods of experimenter-defined central and peripheral items (Dalton & Daneman, 2006; Wright & Stroud, 1998). By our definition, central items are necessarily easier to remember than peripheral items, and would likely be more resistant to misinformation (Holliday, Douglas, &

Hayes, 1999; Marche, 1999; Okado & Stark, 2005). Therefore, we expect the central items to be accompanied by a much smaller misinformation effect than the peripheral items. The interesting unknown is whether initial testing would increase suggestibility for these central items.

Importantly, even if item centrality were found to affect the magnitude of RES (e.g., that only peripheral details are susceptible to RES), the phenomenon would not be rendered trivial for several reasons. First, event details that are less memorable or less central are not necessarily less important in the context of a criminal investigation. Crimes are often complex and what constitutes important information cannot be easily predicted. Indeed, the course of an investigation can often lead investigators to reassess the relative importance or unimportance of specific pieces of evidence. Second, and perhaps even more importantly, witnesses who incorporate peripheral details into their memory reports are perceived as more credible than those who omit such details (Bell & Loftus, 1988; 1989; Wells & Leippe, 1981). Thus, examining potential moderators of RES (such as item centrality) could help to further illuminate its application in legal contexts.

Cognitive Mechanisms Underlying Retrieval-Enhanced Suggestibility

Proposed theoretical mechanisms underlying RES are beginning to emerge.[1] One

---

[1] One account suggests that memory for the original event must undergo reconsolidation after initial retrieval (Hardt, Einarsson, & Nader, 2010), and presentation of the misinformation disrupts this reconsolidation process, thus impairing later retrieval of the original memory (Chan et al., 2009; Chan et al., 2012; Chan & LaPaglia, 2013). We do not dwell on the reconsolidation hypothesis in this paper because these experiments were not designed to address this possibility.

account ascribes RES to preferential encoding of the misinformation following initial testing (Chan et al., 2009; Thomas, Bulevich, & Chan, 2010). For explication purposes, we refer to this notion as the *attention allocation* hypothesis (Chan, Wilford, & Hughes, 2012; Thomas et al., 2010). The attention allocation account posits that the initial test questions serve as cues for the later presented misinformation. For example, if a participant is asked about the type of vehicle driven by the perpetrator, this question might draw attention to that detail later, causing participants to preferentially encode the relevant misinformation. Essentially, this is similar to the prequestions effect in the education literature, where providing students with a set of questions before they read an essay often facilitates learning of that essay. More specifically, the prequestions effect is characterized by enhanced recall of information specifically targeted by or related to the prequestions, whereas recall of information unrelated to the prequestions is diminished (Hamaker, 1986). The occurrence of this effect has been attributed to increased attention paid to the materials queried by the prequestions at the expense of other materials (Lewis & Mensink, 2012). When this experimental phenomenon is applied to the present context, RES can be explained by increased attention to the later misinformation because the initial test included questions related to the misinformation. To be clear, as is presently conceived, the attention allocation hypothesis suggests that RES is the direct result of asking witnesses specific questions during the initial test.

In the present experiments, we provide a test for an extension of the attention allocation account by administering the initial test in a free recall format. All previous demonstrations of RES have been limited to protocols with an initial cued recall test. It is therefore unclear how initial free recall would affect encoding of subsequent misinformation and whether RES would occur in this context. It is well known that free recall differs from cued recall on a number of dimensions, including a heavier reliance on recollection (Hamilton & Rajaram, 2003; Tulving, 1985), self-generated retrieval cues (Burgess & Shallice, 1996), and organization processes (Zaromb & Roediger, 2010). Although a considerable amount of research has demonstrated that testing can potentiate subsequent learning (e.g., Arnold & McDermott, 2013; Izawa, 1970), far less is known about the processes by which free recall affects subsequent learning of complex prose materials (Dunlosky, Rawson, & McDonald, 2009).

In the present experiments, misinformation was presented in an audio narrative that recapitulated much of the witnessed event. Prior free recall may lead to better organization and increased accessibility of details of the witnessed event, which may in turn facilitate comprehension and encoding of the postevent narrative (Arnold & McDermott, 2013; Wissman, Rawson, & Pyc, 2011). It is also possible that by completing the initial free recall test, participants are better able to anticipate the level of details required to perform well on any subsequent memory tests, thus causing a shift in encoding strategy towards remembering details (Britton, Glynn, Muth, & Penland, 1985; McCrudden & Schraw, 2007). Finally, performing an initial free recall test provides participants with information about what they have learned well and what they have not (Lachman & Laughery, 1968; Padilla-Walker & Poole, 2002; Thompson, Wenger, and Bartling, 1978). When a relearning opportunity arises, the initial test allows participants to better allocate their attentional resources to the items that were not well learned (Battig, Allen, & Jensen, 1965; Son & Kornell, 2008). Of particular relevance to the present study is whether prior free recall would have different effects on the central and peripheral misinformation. Because peripheral details are unlikely to be reported during the initial free recall test (and indeed this was how we operationally define peripheral items in the present study), the tested participants might be particularly “sensitive” to these details when they appear during the postevent narrative (by virtue of their noticing that these details were omitted from their free recall report earlier), thus leading to an RES effect for the peripheral details.

As this brief theoretical analysis demonstrates, there are reasons to believe that RES would occur following initial free recall. We caution, however, that this is not a foregone conclusion and the attention allocation hypothesis would have to be broadened substantially from its present form to account for a free recall-based RES effect.

Overview of the Present Experiments

In three experiments, we examined whether RES is sensitive to variations in test format and item centrality. In Experiment 1, in keeping with the typical RES procedure, both the initial and final tests were cued recall. This experiment served two purposes. First, we examined whether central and peripheral details are differentially susceptible to RES in the standard procedure. Second, because it is unclear whether RES would occur with a free recall initial test, it is important to first establish the effect with the present materials and participant population.

In Experiment 2, the initial test was administered in a free recall format. This change allowed us to examine whether an initial cued recall test is necessary to produce RES. In Experiment 3, we further tested the generality of RES by using free recall for both the initial and final tests. This experiment serves two purposes. First, it provides an even more stringent test for the generality of RES. RES has been demonstrated across a variety of final test formats including cued recall (Chan et al., 2009), forced choice recognition (Thomas et al., 2010), and source discrimination (Chan et al., 2012), but it has yet to be demonstrated in a free recall final test. More importantly, because free recall is based heavily, though certainly not entirely (McCabe, Roediger, & Karpicke, 2011; Yonelinas & Jacoby, 2012), on recollection (Tulving, 1985), demonstrating RES in a final free recall test would buttress previous claims that misinformation reported in this procedure is due to true source misattributions rather than guessing (Chan et al., 2012).

Experiment 1

Method

Participants. Sixty undergraduate students (*N* = 30 for each testing condition) participated in this experiment for course research credit.

Design. We used a 2 (Test Condition: test, no-test) x 3 (Item Type: central, peripheral, or neutral) mixed-factorial design. Test condition was manipulated between-subjects whereas item type was manipulated within-subjects. Note that because all participants were administered a final test, the test and no-test labels refer only to the presence or absence of the initial test.

Materials and procedure. We conducted a pilot study to determine item centrality. Participants (*N* = 20) viewed the 8 min critical event video of a museum burglary. Afterwards, participants were given 20 min to type into a blank Word document, in as much detail as possible, the events they remembered from the video. Each narrative was coded into idea units for scoring. Idea units were considered to be the smallest component of words or phrases that conveyed a meaningful idea. For instance, any early mention of the burglar “zip-lining” or “sliding down a rope” would be considered one idea unit concerning how the burglar got to the roof of the museum. Six units recalled frequently were defined as central items (*M* = .74, *SD* = .23), and six units not recalled frequently were defined as peripheral items (*M* = .19, *SD* = .12). An additional six units with a range of recall frequencies were designated as neutral (never misled) items (*M* = .34, *SD* = .26). The neutral items were included so that not all questions in the cued recall tests queried details that were contradicted by misinformation. These central, peripheral, and neutral items dealt with different video details and remained consistent across participants and experiments (refer to Appendix).

In Experiment 1, all participants first viewed the same burglary video used in the pilot study. Afterwards, participants in the test condition completed a cued recall test comprising 18 non-leading questions (e.g., When the burglar worked to retrieve the diamond in the video, what apparatus, if any, was used to retrieve the diamond?). Each of the 18 questions queried one of the six previously designated neutral, peripheral, or central items. Participants were provided 30 sec to type in their answers. After each question, participants were asked to provide a confidence rating on a scale from 1 (not very confident) to 5 (very confident). No feedback was provided and participants were not required to answer each question—questions could be left blank but could not be skipped. Following the initial test (which took 10 min), participants in the test condition played a videogame for 5 min as a distractor activity. Participants in the no-test condition played the same videogame for 15 min.

All participants then listened to an audio narrative (~6 min) via headphones. Participants were told that the narrative recapped the burglary video, but they were not told anything specific regarding the veracity of the narrative (or that it included misinformation). Two versions of the narrative were constructed and counterbalanced across participants. The two narratives were identical with the exception of the twelve pieces of misinformation (e.g., Version A: the burglar used two identical black claws with 3-prongs to retrieve the diamond; Version B: the burglar used one silver mechanical 3-pronged arm to retrieve the diamond; Correct Answer: the burglar used two silver metal arms with 2 and 3-prongs to retrieve the diamond, refer to Appendix). These different sets of misinformation were created to increase generalizability. The six neutral items were not mentioned in either version of the narrative (although they were presented in the video). After the audio narrative, all participants completed a final cued recall test that was identical to the initial test taken by participants in the test condition. Participants were specifically instructed to answer all questions based on what they remembered from the video.

Results and Discussion

We focused our data analyses on the central and peripheral misinformation items, but for the sake of completeness, data for the neutral items are also presented (refer to top of Table 1). Responses in the cued recall test were coded into one of four categories: *correct*, *misinformation*, *other intrusions*, or *no answer* (i.e., blank or “I don’t know”). Other intrusions included all responses that matched neither the misinformation nor the correct answer. A second coder re-coded 25% of the initial and final cued recall data to establish coding consistency—interrater agreement (see Lane, Mather, Villa, & Morita, 2001) reached a proportion of .83 and .87 for the initial and final test, respectively.

Initial cued recall test. Initial test data will not be discussed in detail but are displayed at the top of Table 2. Of most relevance to our purposes, spontaneous misinformation generation was extremely low for the central items (.01) and fairly low for the peripheral items (.11). Spontaneous misinformation generation refers to the proportion of misinformation responses reported before the misinformation was introduced via the audio narrative (i.e., misinformation responses that were spontaneously reported without suggestions). We now report results from the final test.

Final cued recall test. Confidence data for the final test are reported in Table 1 (for neutral items) and Table 3 (for the central and peripheral items). We do not elaborate on the data regarding confidence because they were not influenced by initial testing.

Initial testing increased reports of misinformation on the final test for peripheral items – an RES effect, $t(58) = 2.38$, $d = .63$, $p < .03$ (.52 for test and .36 for no-test, see Figure 1). Tested individuals were also less likely to report the correct answers for the peripheral items ($M = .21$) than the non-tested individuals ($M = .32$), $t(58) = 2.35$, $d = .62$, $p < .03$. Note that correct and misinformation recall probabilities were not entirely complimentary due to the presence of the *other intrusions* and *no answer* response categories. A conditional analysis was performed to examine whether initial recall accuracy affected subsequent suggestibility. Not surprisingly, the effect of misinformation (as measured by the proportion of misinformation recall on the final test) was significantly lower for items that participants correctly recalled during the initial test ($M = .44$) relative to items that participants were not able to correctly recall during the initial test ($M = .58$), $t(26) = 2.50$, $d = .47$, $p =.02$. Note that we define "not able to recall correctly" here as all incorrect responses, including "other intrusions," "no answer," and "spontaneous false recall." More interestingly, the misinformation recall probability for these initially correct items was still numerically (though not significantly, $M = .44$) higher than that of the non-tested participants ($M = .36$), $t(26) = .97$, $d = .38$, $p =.34$. This result is particularly impressive given the inherent item selection differences built into such a comparison (i.e., the initially correct items for the tested participants necessarily reflected the easier and more memorable ones, whereas both easy and difficult items were included for the non-tested participants).

Testing did not affect the likelihood that one would report misinformation for the central items, $t(58) = 1.41$, $d = .37$, $p = .17$. In fact, recall probability of the central misinformation was close to floor, although the difference was in the direction of RES ($M = .11$ for test and $M = .06$ for no-test). Moreover, testing did not alter correct recall probability of central items ($M = .54$ for both test and no-test), $t(58) = .091$, $d = .02$, $p = .93$. Given the way we defined central misinformation (as items with high free recall probabilities), it is not too surprising that these items were more resistant to misinformation than the peripheral items. Moreover, unlike the peripheral items, participants' susceptibility to later central misinformation was not significantly affected by whether one correctly recalled the item ($M = .14$) or not ($M = .07$) during the initial test, $t(25) = 1.48$, $d = .28$, $p = .15$.

In sum, we generalized the RES effect to a new set of materials, though the effect appeared to be restricted to peripheral details. We provide a more thorough discussion regarding item centrality in the General Discussion.

Experiment 2

In Experiment 2, we examined the influence of completing an initial free recall test on subsequent eyewitness suggestibility. As stated in the Introduction, free recall is a major component of several investigative interview techniques and 911 protocols, and yet, all extant demonstrations of RES occurred with a cued recall initial test. Therefore, administering the initial retrieval attempt in free recall provides an important test for the generality of RES. Theoretically, according to the attention allocation hypothesis, initial testing exacerbates eyewitness suggestibility because it inadvertently enhances encoding of the later presented misinformation. Again, as is presently conceived, the attention allocation hypothesis makes no explicit predictions about whether initial free recall would lead to RES. However, a modified version of the hypothesis, which takes into account the literature on test-potentiated learning, can lead one to predict an RES effect following initial free recall.

Method

Participants. One hundred and twenty undergraduate students ($N = 60$ for each testing condition) participated in this experiment.

Materials and Procedure. All materials for Experiment 2 were identical to Experiment 1. The protocol for Experiment 2 differed from the previous experiment in three ways. All of the protocol changes served to provide a better examination of whether initial free recall produces RES. This objective differed from Experiment 1, which primarily served as a replication of the focal RES effect with the present materials and participant population. First, participants in the test condition were given a 20 min free recall test (as opposed to a 10 min cued recall test). The free recall test required more time than the cued recall test because participants needed to recall event details covering the entire burglary in the former, but only details queried by the specific questions in the latter. Participants were instructed to write a detailed description of the video clip they witnessed, and to type their response into a blank Word document. They were encouraged to utilize the entire 20 min period and to be as detailed as possible (including facts about actions, settings, props, etc.). Moreover, at the end of each 5 min interval, participants were told to "draw" a horizontal line on their Word document. This was done to examine whether participants used (or needed) the entire 20 min for recall. Instead of recalling the video event, participants in the no-test condition were given 20 min to type out a description of a childhood story they had seen visually depicted (e.g., Snow White and the Seven Dwarfs); this distractor task served to better equate the level of verbal processing required between the test and no-test conditions. The second procedural difference required that all participants watched a ~22 min distractor video after typing out their narratives (rather than playing Tetris). Participants then listened to the same post-event audio misinformation narrative. Third, an additional distractor task was inserted between the audio narrative and final test. In this task, participants counted backwards by three starting at 500 for five min. The final test was still administered in a cued recall format and matched that in Experiment 1.

Results and Discussion

Initial free recall test. A single coder counted the frequency of central, peripheral, and neutral items mentioned in the initial free recall narrative. All of these items were then coded as *correct*, *misinformation*, or *other intrusions*. Any critical items excluded from the free recall narratives (i.e., not counted as correct, misinformation, or other intrusions) were considered omissions or *no answer* responses. Another coder coded 25% of the narratives, and interrater agreement was high (.90). In the initial free recall test, participants accurately reported .27 of the neutral items (refer to middle of Table 1), .72 of the central items, and .21 of the peripheral items (see middle of Table 2). These data were similar to those from the pilot experiment. Further, an analysis of the mean number of critical items (i.e., the 18 items designated as central, peripheral, and neutral) recalled at each 5 min interval reveals that participants did utilize all 20 min in the free recall task (refer to top of Table 4). Spontaneous misinformation generation was rare for both central items (M = .02) and peripheral items (M = .04).

Final cued recall test. A single coder coded responses in the recall test into one of four categories: *correct*, *misinformation*, *other intrusions*, or *no answer* (i.e., blank or "I don't know"). A different coder re-coded 25% of the final cued recall responses to verify the consistency of the coding—the coders matched at a proportion of .86. Once again, confidence ratings were not affected by initial testing and these data are presented in Table 1 (neutral items) and Table 3 (central and peripheral items).

Most important for present purposes is that taking an initial free recall test increased the report of misinformation in the final cued recall test (refer to Figure 2). Moreover, similar to Experiment 1, this RES effect was evident for peripheral misinformation only. Initial free recall testing increased misinformation recall probability (M = .39 for test and M = .23 for no-test) during the final cued recall test for peripheral items, $t(118) = 3.75$, $d = .69$, $p < .001$. Interestingly, the size of this RES effect is comparable to the one reported in Experiment 1 ($d = .63$), which suggests that altering the initial test to free recall did not weaken the impact of initial testing on subsequent eyewitness suggestibility. Tested participants also recalled fewer peripheral items correctly relative to non-tested participants ($M = .22$ for test and $M = .29$ for no-test), $t(118) = 2.01$, $d = .37$, $p < .05$. Thus, we provided an important extension of the RES finding to protocols with an initial free recall test.

Similar to Experiment 1, central misinformation produced very little false recall regardless of whether participants were tested previously ($M = .08$ for test versus $M = .04$ for no-test), $t(118) = 1.61$, $d = .30$, $p = .11$. Correct recall probability for central items also did not differ based on testing condition ($M = .55$ for test and $M = .53$ for no-test), $t < 1$, $p = .70$.[2]

In sum, we found that RES is preserved with an initial free recall test. These results show that a cued recall initial test is not necessary for RES to occur. We discuss the theoretical and applied implications of this finding following presentation of the results from Experiment 3.

---

[2] We opted not to conduct a conditional analysis of final false recall probability based on initial recall accuracy for Experiments 2 and 3. There were too few initially incorrect central items (1.7 observations per participant in Experiment 2 and 0.9 items per participant in Experiment 3) and initially correct peripheral items (1.26 observations per participant in Experiments 2 and 3, see Table 2) to produce reliable, stable results for the comparison. Note that "incorrect" central items include those coded as "incorrect", "spontaneous misinformation", and omitted items.

Experiment 3

In Experiment 3, we examined whether the RES effect generated by an initial free recall test would occur when the final test was also administered in a free recall format. This is important because details reported spontaneously (as in free recall) are often perceived to be more accurate than those reported following prompts (as in cued recall). For example, in the trainers' manual for law enforcement issued by the Department of Justice (2003), law enforcement officials are instructed specifically to encourage the witness to volunteer information without prompting, because "unprompted responses tend to be more accurate than those given in response to an interviewer's questioning" (p. 19). Consequently, if the RES effect occurs with a final free recall test, it would further bolster the idea that RES has important legal implications. Further, despite the vast literature on eyewitness suggestibility, a surprisingly small number of studies have examined the influence of misinformation on free recall performance. Of the few studies that used free recall as the final test, some have failed to demonstrate any discernable misinformation effect (Bjorklund, Bjorklund, Brown, & Cassel, 1998; Bjorklund, Cassell, Bjorklund, et al., 2000). Therefore, whether initial testing can produce confidently held false memories for misinformation that persist in a free recall final test remains to be seen (cf., Roediger & McDermott, 1995).

Method

Participants. One hundred and twenty students participated in this experiment, with 60 each in the test and no-test conditions.

Materials and procedure. All experimental protocols were identical to Experiment 2 with the following exceptions. First, the distractor video was shortened to ~18 min (from 22 min) to allow more time for the final free recall test (as opposed to the shorter cued recall final test used in Experiments 1 and 2). Both the initial and final tests were free recall—the initial test was 20 min (just as in Experiment 1), but the final test was increased to 25 min. The final test was increased to 25 min in order to determine the time at which recall approached asymptote. The instructions for the final recall test were identical to those for the initial free recall test. In short, participants watched the event video, completed a 20 min free recall test or recalled a childhood story, watched an 18 min distractor video, listened to the misinformation audio, performed a 5 min distractor task, and concluded with a 25 min final free recall test.

Results and Discussion

A single coder counted the frequency of central, peripheral, and neutral items mentioned in both the initial and final free recall narratives. All of these items were then coded as *correct*, *misinformation*, or *other intrusions*, and omitted critical details were coded as "no answer" responses. A second coder independently coded 25% of the narratives. Interrater agreement reached a proportion of .93 and .83 for the initial and final test, respectively.

Initial free recall test. Participants accurately recalled .29 of the neutral items (refer to bottom of Table 1), .85 of the central items, and .21 of the peripheral items (see the bottom of Table 2). The majority of participants again appeared to engage in the free recall task for the entire 20 min (see middle of Table 4). The frequency of spontaneous misinformation generation was again very low for central (M = .01) and peripheral items (M = .03). Overall, as expected, these data are similar to those in Experiment 2.

Final free recall test. Most importantly, the RES effect persisted for peripheral items even when both the initial and final tests were administered in a free recall format. Similar to previous reports, the free recall format is highly resistant to the influence of misinformation (Bjorklund et al., 1998, 2000), such that the non-tested participants recalled very little of the peripheral misinformation they encountered (*M* = .10). Remarkably, initial testing nearly doubled false recall probability for these items (*M* = .19), $t(118) = 2.57$, $d = .47$, $p < .02$ (refer to Figure 3). False recall of the central items remained low, and no significant RES effect was detected (*M* = .05 for test and *M* = .03 for no-test), $t(118) = 1.16$, $d = .22$, $p < .25$. Moreover, testing did not affect correct recall probability of the central items (*M* = .86 for test and *M* = .88 for no-test), $t < 1$, $p = .64$. When recall probability of the critical items were examined based on time (see bottom of Table 4), it is clear that some participants used the full 25 min for the free recall task, although performance appears to begin reaching asymptote during the 20th and 25th minute.

General Discussion

Three primary findings emerged from these experiments. First, initial retrieval exacerbated suggestibility regardless of whether the initial test format was cued recall (Experiment 1) or free recall (Experiments 2 and 3). Second, RES occurred regardless of whether the final test was cued recall (Experiments 1 and 2) or free recall (Experiment 3). Third, the magnitude of the RES effect was greater for peripheral details than for central details. These results extend the implications of RES to eyewitness suggestibility in legal contexts. Before discussing these implications, we first address whether central details are susceptible to the influence of RES.

Are Central Details Immune to Retrieval-Enhanced Suggestibility?

In all three experiments, no significant RES effect was found for the central details. In fact, it appears that the central details were generally immune to misinformation, as recall probabilities of the central misinformation were consistently near floor. Although one might be tempted to conclude that central items in a witnessed event represent a boundary for RES, a closer examination suggests a more complex story. In every experiment, initial testing increased, albeit not significantly, the recall probability of the central misinformation items (.06 versus .11 in Experiment 1, .04 versus .08 in Experiment 2, and .03 versus .05 in Experiment 3). To further scrutinize the influence of initial testing on central misinformation, we increased statistical power by pooling the data across all three experiments into a meta-analysis. The result from this analysis shows that initial testing significantly increased false recall of the central items (.04 for no-test versus .07 for test), overall d = .25, .95 CI of d =.03 to .48. Admittedly, this effect is small and required the combined power of three experiments to reach statistical significance. However, given the difficulty in obtaining a sizable misinformation effect with central items (e.g., Heath & Erickson, 1998; Roebers & McConkey, 2003; Sutherland & Hayne, 2001), it is quite impressive that initial testing nearly doubled the misinformation effect for these items. This finding suggests that central details of a witnessed event may not be immune to the influence of RES.

Applied Implications

Eyewitnesses are often asked to recount the details of their experience to authorities on multiple occasions (e.g., 911 operators, police officers, lawyers, jurors, etc.). Further, these recall attempts represent a small portion of the times a witness will inevitably recall or think about a witnessed event. Any event warranting an eyewitness account is likely of enough significance for most people to share with friends and family. These events are also likely to be mentally reviewed spontaneously by the eyewitness. Such informal retrieval attempts are of particular relevance to the present research because they conform more closely to free recall (rather than cued recall) protocols. Moreover, recent evidence has shown that people spontaneously recall past events far more often than was previously realized (Hintzman, 2011;

Rasmussen & Bernsten, 2009). Thus, the extension of RES to situations employing either cued or free recall constitutes a crucial piece of evidence in the generality of the RES phenomenon and its potential implications for legal investigations. The present results might appear to contradict those showing that performing an initial Cognitive Interview can reduce subsequent eyewitness suggestibility (Gabbert, Hope, Fisher, & Jamieson, 2012; Memon, Zaragoza, Clifford, & Kidd, 2010). We caution that significant methodological differences exist between these studies and ours, which makes direct comparisons difficult (e.g., Memon et al. asked participants to fabricate their own misinformation). Moreover, whether completing an initial Cognitive Interview always reduces people's susceptibility to subsequent misleading suggestions is far from clear at this point, as others have reported either no reduction (Roos af Hjelmsäter, Strömwall, & Granhag, 2012) or RES (LaPaglia, Wilford, Rivard, Chan, & Fisher, 2013).

From an applied perspective, perhaps the most important finding here was that RES occurred when both the initial and final tests were free recall. This finding is particularly notable for several reasons: First, most research on the misinformation effect has used either cued recall or recognition as the final test format. This methodological decision stems from the fact that researchers are interested in the effects of misinformation on a particular memory detail. Asking a question about that particular detail naturally provides the best way to assess the effect of misinformation on memory for that detail. This is similar to teachers asking specific questions on an exam, because they want to know if students learned the critical, or important, facts. The problem with this approach in the context of eyewitness memory is that cued recall cannot assess whether an eyewitness would report misinformation spontaneously (i.e., without direct cues to the misinformation-relevant detail). In legal settings, it is often impossible to know if a person has encountered exactly what, if any, misinformation. Consequently, it is not possible to ask specific questions pertaining directly to any misinformation. If misinformation is not spontaneously reported without targeted prompts (as in, e.g., a free recall test), then its influence on memory might be more limited in real-world settings than is suggested by research in the laboratory. In Experiment 3, we showed that participants could be induced to report misinformation in a free recall test (at least for the peripheral details). Even more importantly, initial testing almost doubled the frequency of reported misinformation in free recall.

Theoretical Implications

The data from Experiments 2 and 3 show that completing an initial free recall test can increase people's susceptibility to later presented misinformation. From a theoretical perspective, these data can be accommodated by a modified, broader version of the attention allocation account as outlined in the Introduction, which posits that RES occurs because initial free recall enhances the learning of subsequent misinformation. Thus, it seems that initial testing, regardless of format, alters how people encode subsequent information (de Winstanley & Bjork, 2004; Izawa, 1970; Karpicke, 2009; Son & Kornell, 2008; Thompson, Wenger, & Bartling, 1978).

Although the results are similar across experiments (i.e., they produced a RES effect of similar magnitudes), initial cued recall and free recall might have enhanced subsequent encoding of the misinformation based on different mechanisms. When participants were administered a cued recall initial test, their attention was drawn to the subsequent misinformation because the initial test questions foreshadowed the importance of these specific details. That is, to some extent, participants' allocation of attention was manipulated by the experimenter; just as prequestions can serve to direct students' attention to particularly important information (Hamaker, 1986). In contrast, because no specific questions were asked in the free recall test, any facilitation of misinformation learning during the postevent narrative phase cannot be attributed to experimenter-induced attention to the misinformation. Instead, the free recall initial test likely provided participants with an opportunity to determine what information they could and could not recall (Padilla-Walker & Poole, 2002). When participants listened to the misinformation narrative later, they might notice details that were omitted from (most likely for the peripheral misinformation), or were inconsistent with (likely for the central items), their initial free recall report and thus allocated more effort to encode these details.

Consistent with this notion, people typically spend more time to encode information that they missed on a previous test in multi-trial learning environments (Battig et al., 1965; Son & Kornell, 2008; and more broadly, Nelson, Dunlosky, Graf, & Narens, 1994; Son & Metcalfe, 2000; Thiede & Dunlosky, 1999). Unlike these studies, where the learner has free control over study time, participants had no control over study time in the present study. Because the misinformation was presented via a fixed paced audio narrative, all participants received the same amount of time to process each piece of misinformation. As a result, the increased learning of misinformation in the present study must be attributed to more efficient encoding processes following testing (see also Karpicke, 2009). This notion is related to the theoretical account proposed by Arnold & McDermott (2013), which posited that free recall potentiated subsequent learning due to enhanced organization of the already encoded information. By asking participants to freely recall learned material, they are better able to organize their memory of that learned material. This enhanced organization can lead to better processing of any new or unlearned information by increasing one's awareness that the information is new or unlearned—prioritizing cognitive resources to anything not incorporated into one's current memory.

Consequently, when a second study opportunity is presented (i.e., the audio narrative), the new (mis)information is prioritized and encoded more efficiently.

The enhanced encoding of misinformation could also be attributed to activation of related information during initial testing (Carpenter, 2011; Chan, McDermott, & Roediger, 2006; Grimaldi & Karpicke, 2012). That is, during the initial recall phase, information recalled explicitly and related information receives some activation, and the activation of this information may facilitate subsequent encoding during the misinformation learning phase (i.e., a form of long-term semantic priming that is augmented by strategic, controlled processes, Balota, Black, & Cheney, 1992; Becker, Moscovitch, Behrmann, & Joordens, 1997; Chan et al., 2006). Note that none of the discussed accounts are mutually exclusive and that a constellation of factors could be contributing to RES.

Conclusions

Our findings show that initial recall can increase subsequent eyewitness suggestibility regardless of the recall test format. Contrary to the idea that RES is contingent on an initial cued recall test that draws attention to the later misinformation, the present results indicate that the RES effect might be more pervasive than previously thought. Overall, the effect has been found in free recall, cued recall, recognition (Chan & LaPaglia, 2013; Thomas et al., 2010), and source discrimination (Chan et al., 2012); it occurs at delays of 30 min, 48 hr, and a week between the witnessed event and final test (regardless of whether the misinformation occurs before or after the delay, Chan & Langley, 2011; Chan & LaPaglia, 2011), and it is found in both between- and within-subjects comparisons (Chan & LaPaglia, 2011).

Although RES appears to be robust, it is, fortunately, not universal. First, initial testing can enhance subsequent eyewitness memory performance in the absence of misinformation.

Moreover, a number of studies have found that initial testing can enhance, instead of impair, memory performance even in the face of misinformation (Pansky & Tenenboim, 2010; Thomas et al., 2010). It would be fruitful for future research to attempt to pinpoint when repeated retrieval can exacerbate and when it can inoculate against eyewitness suggestibility (LaPaglia & Chan, 2013). Granted, providing answers to such questions will require a series of systematic and careful investigations. However, given the prevalence and frequency that eyewitnesses are asked to retrieve information in the process of a criminal investigation, a more thorough understanding of the effects of retrieval on suggestibility would be instrumental to devising techniques and policies to enhance the accuracy of eyewitness evidence.

The present research primarily serves to demonstrate the generality of RES. The extension of RES to protocols using free recall poses important implications for legal contexts. RES also adds an interesting dimension to the literature on test-potentiated learning or the interim test effect by demonstrating that the improvements in long-term recall might extend to both true and false information. Broadly speaking, at its core, RES contributes to a long line of research that continues to further our understanding of retrieval processes in human memory (Schacter, Eich, & Tulving, 1978).

References

Arnold, K. M., & McDermott, K. B. (2013). Free recall enhances subsequent learning. *Psychonomic Bulletin & Review*. doi:10.3758/s13423-012-0370-3

Balota, D. A., Black, S. R., & Cheney, M. (1992). Automatic and attentional priming in young and older adults-Reevaluation of the two-process model. *Journal of Experimental Psychology: Human Perception and Performance, 18*(2), 485-502.

Battig, W. F., Allen, M., & Jensen, A. R. (1965). Priority of free recall of newly learned items. *Journal of Verbal Learning and Verbal Behavior*, *4*(3), 175–179. doi:10.1016/S0022-5371(65)80017-2

Becker, S., Moscovitch, M., Behrmann, M., Joordens, S. (1997). Long-term semantic priming: A computation account and empirical evidence. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23*(5), 1059-1082.

Bell, B. E., & Loftus, E. F. (1988). Degree of Detail of Eyewitness Testimony and Mock Juror Judgments. *Journal of Applied Social Psychology*, *18*(14), 1171–1192. doi:10.1111/j.1559-1816.1988.tb01200.x

Bell, B. E., & Loftus, E. F. (1989). Trivial persuasion in the courtroom: The power of (a few) minor details. *Journal of Personality and Social Psychology, 56(5),* 669-679.

Bjorklund, D. F., Bjorklund, B. R., Brown, R. D., & Cassel, W. S. (1998). Children's susceptibility to repeated questions: How misinformation changes children's answers and their minds. *Applied Developmental Science*, *2(2)*, 99-111.

Bjorklund, D. F., Cassell, W. S., Bjorklund, B. R., Brown, R. D., Park C. L., Ernst, K., & Owen, F. A. (2000). Social demand characteristics in children's and adults' eyewitness memory and suggestibility: The effect of different interviewers on free recall and recognition. *Applied Cognitive Psychology*, *14*, 421-433.

Britton, B., Glynn, S., Muth, K. D., & Penland, M. J. (1985). Instructional objectives in text: Managing the reader's attention. *Journal of Literacy Research*, *17*(2), 101–113. doi:10.1080/10862968509547533

Burgess, P. W., & Shallice, T. (1996). Confabulation and the control of recollection. *Memory*, *4*, 359–411.

Carpenter, S. K. (2011). Semantic information activated during retrieval contributes to later retention: Support for the mediator effectiveness hypothesis of the testing effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 37*, 1547-1552.

Chan, J. C. K., & Langley, M. (2011). Paradoxical effects of testing: Retrieval enhances both accurate recall and suggestibility in eyewitnesses. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 37*, 1234-1243.

Chan, J. C. K., & LaPaglia, J. A. (2011). The dark side of testing memory: Repeated retrieval can enhance eyewitness suggestibility. *Journal of Experimental Psychology: Applied*, *17(4)*, 418-432.

Chan, J. C. K., & LaPaglia, J. A. (2013). Impairing existing declarative memory in humans by disrupting reconsolidation. *Proceedings of the National Academy of Sciences of the United States of America*. doi:10.1073/pnas.1218472110

Chan, J. C. K., McDermott, K. B., & Roediger, H. L. (2006). Retrieval-induced facilitation: Initially nontested material can benefit from prior testing of related material. *Journal of Experimental Psychology: General, 135*, 553-571.

Chan, J. C. K., Thomas, A. K., & Bulevich, J.B. (2009). Recalling a witnessed event increases eyewitness suggestibility: The reversed testing effect. *Psychological Science, 20,* 66-73.

Chan, J. C. K., Wilford, M. M., & Hughes, K. L. (2012). Retrieval can increase or decrease suggestibility depending on how memory is tested: The importance of source complexity. *Journal of Memory and Language*, in press. doi: 10.1016/j.jml.2012.02.006

Chrobak, Q. M., & Zaragoza, M. S. (2013). The misinformation effect: Past research and recent advances. In A. Ridley, F. Gabbert, & D. La Rooy (Eds.). *Suggestibility in legal contexts: Psychological research and forensic implications* (pp. 21-44). London: Wiley Blackwell.

Dalton, A. L., & Daneman, M. (2006). Social suggestibility to central and peripheral misinformation. *Memory, 14(4)*, 486-501.

de Winstanley, P. A., & Bjork, E. L. (2004). Processing strategies and the generation effect: Implications for making a better reader. *Memory and Cognition, 32(6)*, 945-955.

Department of Justice, National Institutes of Justice. (2003). *Eyewitness evidence: a trainer's manual for law enforcement*. Retrieved from https://www.ncjrs.gov/nij/eyewitness/188678.pdf

Dunlosky, J., Rawson, K. A., & McDonald, S. L. (2009). Influence of practice tests on the accuracy of predicting memory performance for paired associates, sentences, and text material. In T. J. Perfect & B. L. Schwartz (Eds.), *Applied Metacognition* (pp. 68–92). Cambridge: Cambridge University Press. doi:10.1017/CBO9780511489976.005

Fisher, R. P., & Geiselman, R. E. (1992). *Memory-enhancing techniques for investigative interviewing: The Cognitive Interview.* Springfield, IL: Thomas.

Gabbert, F., Hope, L., Fisher, R. P., & Jamieson, K. (2012). Protecting against misleading post-event information with a Self-Administered Cognitive Interview. *Applied Cognitive Psychology, 26*, 568-575.

Grimaldi, P. J., & Karpicke, J. D. (2012). When and why do retrieval attempts enhance subsequent encoding? *Memory & Cognition, 40,* 505-513.

Hamaker, C. (1986). The effects of adjunct questions on prose learning. *Review of Educational Research, 56(2)*, 212-242.

Hamilton, M., & Rajaram, S. (2003). States of awareness across multiple memory tasks: Obtaining a "pure" measure of conscious recollection. *Acta Psychologica*, *112*, 43–69.

Hardt, O., Einarsson, E. O., & Nader, Karim. (2010). A bridge over troubled water: Reconsolidation as a link between cognitive and neuroscientific memory research traditions. *Annual Review of Psychology, 61,* 141-167.

Heath, W. P., & Erickson, J. R. (1998). Memory for central and peripheral actions and props after varied post-event presentation. *Legal and Criminological Psychology, 3*, 321-346.

Hintzman, D. L. (2011). Research strategy in the study of memory: Fads, fallacies, and the search for the "coordinates of truth". *Perspectives on Psychological Science, 6(3)*, 253-271.

Holliday, R. E., Douglas, K. M., & Hayes, B. K. (1999). Children's eyewitness suggestibility: Memory trace strength revisited. *Cognitive Development, 14,* 443-462.

Home Office (1992). *Memorandum of good practice on video recorded interviews with child witnesses for criminal proceedings.* London: HMSO.

Home Office (2001). *Achieving best evidence in criminal proceedings: Guidance for vulnerable or intimidated witnesses including children*. London: Communication Directorate.

Izawa, C. (1970). Optimal potentiating effects and forgetting-prevention effects of tests in paired-associate learning. *Journal of Experimental Psychology, 83(2),* 340-344.

Karpicke, J. D. (2009). Metacognitive control and strategy selection: Deciding to practice retrieval during learning. *Journal of Experimental Psychology: General, 183(4)*, 469-486.

Lachman, R., & Laughery, K. R. (1968). Is a test trial a training trial in free recall learning? *Journal of Experimental Psychology*, *76*, 40–50. doi:10.1037/h0025278

Lamb, M. E. (2007). The “approximation rule”: Another proposed reform that misses the target. *Child Development Perspectives*, *1(2)*, 135-136.

Lane, S. M., Mather, M., Villa, D., & Morita, S. (2001). How events are reviewed matters: Effects of varied focus on eyewitness suggestibility. *Memory & Cognition*, *29*, 940–947.

LaPaglia, J. A., & Chan, J. C. K. (2013). *Testing can enhance or reduce suggestibility depending on how misinformation is presented.* Manuscript submitted for publication.

LaPaglia, J. A., Wilford, M. M., Rivard, J. R., Chan, J. C. K., Fisher, R. P. (2013). *Misleading suggestions can alter later memory reports even following a Cognitive Interview.* Manuscript submitted for publication.

Lewis, M. R., & Mensink, M. C. (2012). Prereading questions and online text comprehension. *Discourse Processes, 49:5*, 367-390.

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

Loftus, E. F. (1979). Reactions to blatantly contradictory information. *Memory & Cognition, 7(5)*, 368-374.

Loftus, E. F., & Palmer, J. C. (1974). Reconstruction of automobile destruction: An example of the interaction between language and memory. *Journal of Verbal Learning and Verbal Behavior*, 13, 585-589.

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

Masson, M. E. J., & Loftus, G. R. (2003). Using confidence intervals for graphically based data interpretation. *Canadian Journal of Experimental Psychology*, *57(3)*, 203-220.

McCabe, D. P., Roediger, H. L., III, & Karpicke, J. D. (2011). Automatic processing influences free recall: converging evidence from the process dissociation procedure and remember-know judgments. *Memory & Cognition*, *39*(3), 389–402.

McCrudden, M. T., & Schraw, G. (2007). Relevance and Goal-Focusing in Text Processing. *Educational Psychology Review*, *19*(2), 113–139. doi:10.1007/s10648-006-9010-7

Memon, A., Zaragoza, M., Clifford, B. R., & Kidd, L. (2010b). Inoculation or antidote? The effects of cognitive interview timing on false memory for forcibly fabricated events. *Law and Human Behavior, 34,* 105-117.

Nelson, T. O., Dunlosky, J., Graf, A., & Narens, L. (1994). Utilization of metacognitive judgments in the allocation of study during multitrial learning. *Psychological Science*, *5*, 207–213.

Okado, Y. & Stark, C. E. L. (2005). Neural activity during encoding predicts false memories created by misinformation. *Learning & Memory, 12,* 3-11.

Padilla-Walker, L. M., & Poole, D. A. (2002). Memory for previous recall: A comparison of free and cued recall. *Applied Cognitive Psychology, 16(5)*, 515-524. doi: 10.1002/acp.809

Pansky, A., & Tenenboim, E. (2010). Inoculating against eyewitness suggestibility via interpolated verbatim versus gist testing. *Memory & Cognition*, *39*, 155-170.

Paz-Alonso, P. M., & Goodman, G. S. (2008). Trauma and memory: Effects of post-event misinformation, retrieval order, and retention interval. *Memory, 16(1),* 58-75.

Porter, S., Spencer, L., & Birt, A. R. (2003). Blinded by emotion? Effect of the emotionality of a scene on susceptibility to false memories. *Canadian Journal of Behavioural Science, 35,* 165-175.

Rasmussen, A. S., & Bernsten, D. (2009). The possible functions of involuntary autobiographical memories. *Applied Cognitive Psychology, 23(8)*, 1137-1152.

Roebers, C. M. & McConkey, K. M. (2003). Mental reinstatement of the misinformation context and the misinformation effect in children and adults. *Applied Cognitive Psychology, 17,* 477-493.

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(4)*, 803-814.

Roediger, H. L., Putnam, A. L., & Smith, M. A. (2011). Ten benefits of testing and their applications to educational practice. *Psychology of Learning and Motivation*, *55*, 1-32.

Roos af Hjelmsäter, E. R., Strömwall, L. A., & Granhag, P. A. (2012). The Self-Administered Interview: a means of improving children's eyewitness performance? *Psychology, Crime and Law*, *18*(10), 897–911. doi:10.1080/1068316X.2011.582844

Saunders, J. (2009). *Memory impairment in the weapon focus effect. Memory & Cognition, 37(3), 326-335.*

Schacter, D. L., Eich, J. E., & Tulving, E. (1978). Richard Semon's theory of memory. *Journal of Verbal Learning and Verbal Behavior*, *17*(6), 721–743. doi:10.1016/S0022-5371(78)90443-7

Son, L. K., & Kornell, N. (2008). Research on the allocation of study time: Key studies from 1890 to the present (and beyond). In J. Dunlosky & R. A. Bjork (Eds.), *Handbook of metamemory and memory* (pp. 333-351). New York: Psychology Press.

Son, L. K., & Metcalfe, J. (2000). Metacognitive and control strategies in study-time allocation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*(1), 204. doi:10.1037//0278-7393,26.1.204

Sutherland, R., & Hayne, H. (2001). Age-related changes in the misinformation effect. *Journal of Experimental Child Psychology*, *79*, 388-404.

Thiede, K. W., & Dunlosky, J. (1999). Toward a general model of self-regulated study: An analysis of selection of items for study and self-paced study time. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*, 1024–1037.

Thomas, A. K., Bulevich, J. B., & Chan, J. C. K. (2010). Testing promotes eyewitness accuracy with a warning: Implications for retrieval enhanced suggestibility. *Journal of Memory and Language, 63,* 149-157.

Thompson, C. P., Wenger, S. K.,& Bartling, C. A. (1978). How recall facilitates subsequent recall: A reappraisal. *Journal of Experimental Psychology: Human Learning and Memory, 4,* 210-221.

Tulving, E. (1985). Memory and consciousness. *Canadian Psychology*, *26*, 1–12.

Wells, G. L., & Leippe, M. R. (1981). How do triers of fact infer the accuracy of eyewitness identifications? Using memory for peripheral detail can be misleading. *Journal of Applied Psychology*, *66*(6), 682–687. doi:10.1037/0021-9010.66.6.682

Wissman, K. T., Rawson, K. A., Pyc, M. A. (2011). The interim test effect: Testing prior material can facilitate the learning of new material. *Psychonomic Bulletin & Review*, *18*, 1140-1147.

Wright, D. B., & Stroud, J. N. (1998). Memory quality and misinformation for peripheral and central objects. *Legal and Criminological Psychology, 3,* 273-286.

Yonelinas, A. P., & Jacoby, L. L. (2012). The process-dissociation approach two decades later: Convergence, boundary conditions, and new directions. *Memory & Cognition*, *40*(5), 663–680. doi:10.3758/s13421-012-0205-5

Yuille, J. C., Hunter, R., Joffe, R., & Zaparniuk, J. (1993). Interviewing children in sexual abuse cases. In G. S. Goodman & B. L. Bottoms (Eds.), *Child victims, child witnesses: Understanding and improving testimony* (pp. 95-115). New York: Guilford Press.

Zaromb, F. M., & Roediger, H. L. (2010). The testing effect in free recall is associated with enhanced organizational processes. *Memory & Cognition*, *38*(8), 995–1008. doi:10.3758/MC.38.8.995

Table 1

Probability and Confidence of Response for Neutral Items in Experiments 1-3

	Response Type		
	Correct	Other Intrusions	No Answer
Experiment 1			
Initial Cued Recall	.83 (.19)	.12 (.16)	.05 (.11)
Confidence	3.86 (.49)	2.55 (1.06)	1.14 (.38)
Final Cued Recall			
No-test	.83 (.18)	.15 (.15)	.02 (.05)
Confidence	3.84 (.51)	2.34 (1.18)	2.33 (1.53)
Test	.87 (.15)	.10 (.29)	.03 (.06)
Confidence	3.86 (.59)	1.20 (.45)	2.95 (1.19)
Experiment 2			
Initial Free Recall	.27 (.20)	.00 (.00)	.73 (.20)
Final Cued Recall			
No-test	.71 (.27)	.20 (.19)	.09 (.17)
Confidence	3.89 (.80)	2.81 (1.14)	1.50 (.99)
Test	.76 (.21)	.20 (.20)	.04 (.08)
Confidence	3.96 (.77)	1.46 (.88)	2.38 (1.05)
Experiment 3			
Initial Free Recall	.29 (.18)	.03 (.07)	.68 (.18)
Final Free Recall			
No-test	.36 (.21)	.02 (.05)	.63 (.20)
Test	.44 (.22)	.03 (.08)	.54 (.20)

Note. For the free recall tests, "No Answer" responses refer to omissions (i.e., details that participants did not mention in their free recall report). Standard deviations appear in parentheses. Confidence ratings were obtained only for cued recall.

Table 2

Probability of Response and Confidence on Initial Tests As a Function of Item Type and Response Type for Experiments 1-3

	Response Type			
	Correct	Other Intrusions	No Answer	Baserate False Recall
Experiment 1				
Central Items	.62 (.26)	.33 (.23)	.04 (.10)	.01 (.03)
Confidence	4.64 (.78)	4.52 (.72)	2.53 (1.12)	5.00 (.00)
Peripheral Items	.40 (.21)	.40 (.17)	.09 (.12)	.11 (.09)
Confidence	3.61 (.76)	3.09 (1.07)	1.15 (.32)	3.72 (.96)
Experiment 2				
Central Items	.72 (.24)	.00 (.02)	.26 (.24)	.02 (.05)
Peripheral Items	.21 (.15)	.01 (.05)	.74 (.24)	.04 (.20)
Experiment 3				
Central Items	.85 (.17)	.01 (.04)	.14 (.17)	.01 (.03)
Peripheral Items	.21 (.14)	.04 (.08)	.72 (.15)	.03 (.07)

Note. The initial test was cued recall in Experiment 1 and free recall in Experiments 2 and 3. For Experiments 2 and 3, “No Answer” refers to omissions (i.e., details that participants did not mention in their free recall report). Standard deviations appear in parentheses. Confidence ratings were obtained only for cued recall.

Table 3

Mean Confidence on Final Tests As a Function of Item Type and Response Type for Experiments 1-2, Experiment 3 is Excluded Because the Final Test was Free Recall and Therefore Did not Include Confidence Ratings

	Response Type			
	Correct	Other Intrusions	No Answer	Baserate False Recall
Experiment 1				
Final Cued Recall Test				
No-test				
Central Items	4.68 (.39)	4.40 (.73)	3.58 (1.17)	4.50 (.50)
Peripheral Items	3.79 (1.11)	3.38 (1.15)	1.50 (.60)	3.59 (1.11)
Test				
Central Items	4.67 (.40)	4.46 (.54)	N/A	4.47 (.85)
Peripheral Items	3.99 (.70)	3.33 (1.17)	2.40 (1.52)	3.92 (.89)
Experiment 2				
Final Cued Recall Test				
No-test				
Central Items	4.74 (.41)	4.26 (.88)	2.11 (1.58)	3.82 (1.14)
Peripheral Items	3.85 (.98)	3.20 (.95)	1.72 (1.11)	3.51 (.90)
Test				
Central Items	4.64 (.65)	4.27 (.80)	3.21 (1.95)	4.08 (.88)
Peripheral Items	4.02 (.88)	3.31 (.91)	1.32 (.56)	3.57 (1.09)

Note. Standard deviations appear in parentheses.

Table 4

The Mean Cumulative Proportions of Accurately Reported Neutral, Peripheral, and Central Items at Each 5 Min Interval of the Free Recall Tests, Experiment 1 is Excluded Because It Did Not Include Any Free Recall Tests.

	Response Time				
	5 min	10 min	15 min	20 min	25 min
Experiment 2					
Initial Free Recall Test					
Central Items	.15 (.06)	.26 (.13)	.47 (.22)	.73 (.24)	
Peripheral Items	.03 (.07)	.08 (.10)	.15 (.14)	.21 (.15)	
Neutral Items	.04 (.07)	.11 (.13)	.18 (.17)	.28 (.20)	
Experiment 3					
Initial Free Recall Test					
Central Items	.17 (.09)	.39 (.20)	.64 (.24)	.85 (.17)	
Peripheral Items	.02 (.06)	.06 (.09)	.14 (.11)	.21 (.14)	
Neutral Items	.04 (.07)	.12 (.12)	.19 (.14)	.29 (.18)	
Final Free Recall Test					
Central Items	.18 (.09)	.44 (.21)	.71 (.23)	.83 (.18)	.87 (.16)
Peripheral Items	.03 (.07)	.11 (.11)	.16 (.12)	.19 (.13)	.21 (.13)
Neutral Items	.05 (.08)	.15 (.13)	.25 (.17)	.33 (.20)	.40 (.22)

Note. Standard deviations appear in parentheses.

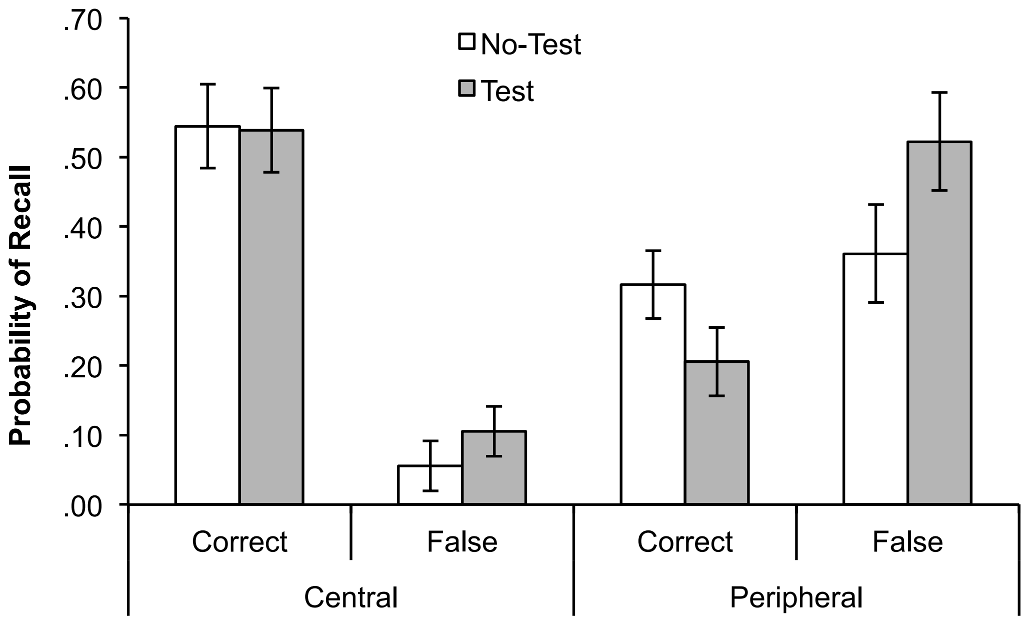

Figure 1. Results from Experiment 1 with an initial and final cued recall test: probability of a correct and false (misinformation) response as a function of item type (central or peripheral), and whether subjects had received an initial cued recall test prior to misinformation exposure via the audio narrative (test versus no-test conditions). The error bars represent the .95 confidence intervals for each between-subjects comparison (Masson & Loftus, 2003).

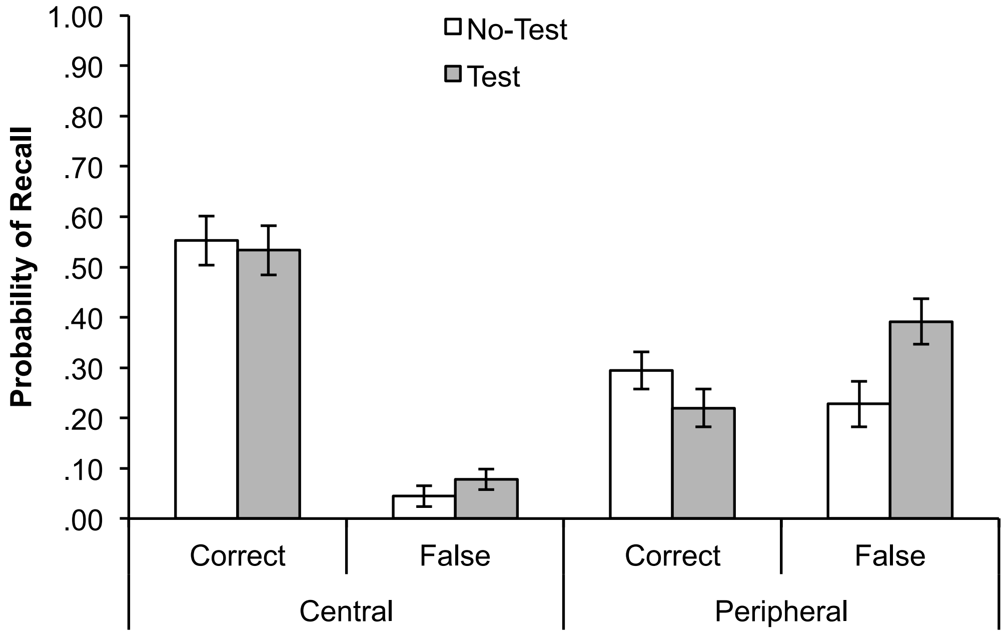

Figure 2. Results from Experiment 2 with an initial free recall test and final cued recall test: probability of a correct and false (misinformation) response as a function of item type (central or peripheral) and whether subjects had received an initial free recall test prior to misinformation exposure via the audio narrative (test versus no-test conditions). The error bars represent the .95 confidence intervals for each between-subjects comparison.

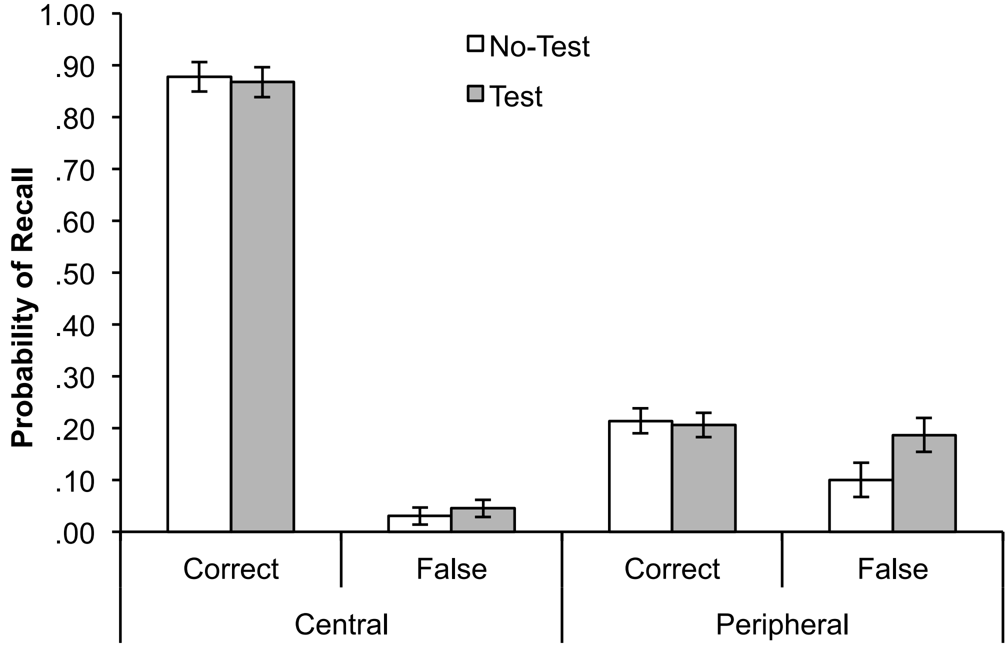

Figure 3. Results from Experiment 3 with an initial and final free recall test: probability of a correct and false (misinformation) response as a function of item type (central or peripheral) and whether subjects had received an initial free recall test prior to misinformation exposure via the audio narrative (test versus no-test conditions). The error bars represent the .95 confidence intervals for each between-subjects comparison.

Appendix

Central Misinformation Items

1. The video began with the burglar on the roof of a building. How did the burglar get from the roof of the building he started on, to the roof of the museum?

   [Correct Answer: Slid down a cable with use of crossbow, Misinformation Version A: Swung on a rope, Misinformation Version B: Jumped]

2. In the video, how did the burglar get under the alarm beams undetected when advancing to the room with the diamond?

   [Correct Answer: Used the crossbow to shoot a rope and pull himself, Misinformation Version A: Used his arms to crawl forward on his belly, Misinformation Version B: Used his legs to slowly propel himself on his back]

3. After getting to the room with the diamond in the video, how did the burglar remove the dome-shaped glass case covering the diamond?

   [Correct Answer: Used the crossbow to shoot a rope and make a pulley, Misinformation Version A: Picked it up carefully with his hands, Misinformation Version B: Used the mechanical arms to remove it]

4. When the burglar worked to retrieve the diamond in the video, what apparatus, if any, is used to retrieve the diamond?

   [Correct Answer: Two silver, metal arms (2 & 3 pronged), Misinformation Version A: Two identical black claws (3-pronged), Misinformation Version B: One silver mechanical arm (3-pronged)]

5. Once the burglar had stolen the diamond, a glove was put in its place, what was the color of the glove that the burglar put in place of the diamond?

[Correct Answer: White, Misinformation Version A: Black, Misinformation Version B: Silver]

6. In the video, how did one of the guards finally discover that there was something going wrong in the museum?

   [Correct Answer: Sees the glove in place of the diamond, Misinformation Version A: Sees the burglar, Misinformation B: Hears the burglar]

Peripheral Misinformation Items

7. As the burglar entered the roof of the museum in the video, the entrance was of a particular shape. What shape was it?

   [Correct Answer: Octagonal, Misinformation Version A: Circular, Misinformation Version B: Square]

8. In the video, upon entering the museum, the burglar found himself in a dimly lit hallway area, what color were the walls painted?

   [Correct Answer: White or Off-White, Misinformation Version A: Beige, Misinformation Version B: Grey]

9. Describe the pedestal, which was used to hold the large baseball-shaped diamond in the big room in the video.

   [Correct Answer: Glass, hourglass-shaped, Misinformation Version A: Marble dolphins, Misinformation Version B: Marble cupids]

10. In the video, a piece of artwork was present on the ceiling of the big room (where the burglar stole the diamond). Describe what the artwork looked like and be as specific as possible (e.g., what patterns were present in the artwork, etc.).

[Correct Answer: Wooden, brown circular patterns, Misinformation Version A: Knights charging, Misinformation Version B: David & Goliath]

11. In the video, the glove had the letter P embellished on it, what color was the letter embellished on the glove?

    [Correct Answer: Silver, Misinformation Version A: White, Misinformation Version B: Gold]

12. In the video, all the museum guards wore hats as part of their uniform, what color were the hats the guards wore?

    [Correct Answer: Dark green, Misinformation Version A: Tan, Misinformation Version B: Dark brown]

Neutral Items

13. In the video, when the two museum guards were shown talking, they were standing in front of what appeared to be a stained glass window. Please describe the design of this window as specifically as possible (e.g., colors, patterns, etc).

    [Correct Answer: Simple shapes, yellow, red, blue]

14. In addition to the stained glass window, what large objects were shown in the video shot of the two guards talking?

    [Correct Answer: Two Greek statues, white, stone]

15. In the video, the dome-shaped glass case that covered the diamond had a handle on it, what shape was the handle of this case?

    [Correct Answer: Moon-shaped, crescent]

16. In the video, a carpet was placed underneath the diamond stand in the large room and surrounded by velvet ropes. Describe the carpet (size, color, pattern, etc.).

[Correct Answer: Red, octagonal]

17. In the video, the alarm is eventually set off, please describe who set the alarm off and how.

    [Correct Answer: The guard accidentally set it off upon seeing the glove]

18. As the burglar escaped from the museum in the video, the guards began to shoot at him, how many guards were shooting at the burglar as he got away?

    [Correct Answer: 3]