1-2009

Recalling a Witnessed Event Increases Eyewitness Suggestibility The Reversed Testing Effect

Jason C.K. Chan
Iowa State University, ckchan@iastate.edu

Ayanna K. Thomas
Tufts University

John B. Bulevich
Rhode Island College

Recalling a Witnessed Event Increases Eyewitness Suggestibility The Reversed Testing Effect

Abstract
People's later memory of an event can be altered by exposure to misinformation about that event. The typical misinformation paradigm, however, does not include a recall test prior to the introduction of misinformation, contrary to what real-life eyewitnesses encounter when they report to a 911 operator or crime-scene officer. Because retrieval is a powerful memory enhancer (the testing effect), recalling a witnessed event prior to receiving misinformation about it should reduce eyewitness suggestibility. We show, however, that immediate cued recall actually exacerbates the later misinformation effect for both younger and older adults. The reversed testing effect we observed was based on two mechanisms: First, immediate cued recall enhanced learning of the misinformation; second, the initially recalled details became particularly susceptible to interference from later misinformation, a finding suggesting that even human episodic memory may undergo a reconsolidation process. These results show that real-life eyewitness memory may be even more susceptible to misinformation than is currently envisioned.

Keywords
False memory, eyewitness memory, misinformation, reconsolidation, testing effect

Disciplines
Cognitive Psychology | Evidence

Comments
This is a manuscript of an article from Psychological Science 20 (2009): 66, doi:10.1111/j.1467-9280.2008.02245.x. Posted with permission.

RESEARCH ARTICLE

Recalling a witnessed event increases eyewitness suggestibility: The reversed testing effect

Jason C.K. Chan
Iowa State University

Ayanna K. Thomas
Tufts University

John B. Bulevich
Rhode Island College

Word Count: 149 in abstract, 3948 in text

Reference Count: 33

Keywords: False memory, eyewitness memory, misinformation, reconsolidation, testing effect

Address correspondence to:
Jason Chan
Department of Psychology
W112 Lagomarcino Hall
Iowa State University
Ames, IA 50011-3180
USA

Email: ckchan@iastate.edu

This is a manuscript of an article from *Psychological Science* 20 (2009): 66, doi:10.1111/j.1467-9280.2008.02245.x. Posted with permission.

Abstract

People's later memory of an event can be altered by exposure to misinformation (Loftus, 1979). The typical misinformation paradigm, however, does not include a recall test prior to the introduction of misinformation, which is what real-life eyewitnesses encounter when they report to a 911 operator or crime scene officer. Because retrieval is a powerful memory enhancer (the testing effect), recalling a witnessed event prior to misinformation should reduce eyewitness suggestibility. Here we show that immediate cued recall actually exacerbates the later misinformation effect for both younger and older adults. This reversed testing effect occurred based on two mechanisms: First, immediate cued recall enhanced learning of the misinformation; second, the initially recalled details became particularly susceptible to interference from later misinformation, suggesting that even human episodic memory may undergo a reconsolidation process. These results show that real-life eyewitness memory may be even more malleable to misinformation than is currently envisioned.

Recalling a witnessed event increases eyewitness suggestibility: The reversed testing effect

Since Bartlett's pioneering work (1932), researchers have repeatedly demonstrated the malleability of human memory (e.g., Jacoby, Woloshyn, & Kelley, 1989; Roediger & McDermott, 1995). Loftus's misinformation paradigm (Loftus, 1979), in particular, has translated traditional verbal learning research findings (e.g., Keppel & Underwood, 1962) into an ecologically relevant setting – namely, that eyewitness memory is malleable to exposure to misinformation. The Loftus paradigm (e.g., Lindsay, Allen, Chan, & Dahl, 2004; Loftus, Miller, & Burns, 1978) includes three phases: First, subjects are exposed to a witnessed event (e.g., a video). Second, subjects are exposed to a narrative description of the witnessed event that contains misinformation. Third, subjects take a memory test for the witnessed event. The typical finding is that people are less likely to recall the correct details following misinformation. Interestingly, this three-phase procedure misses a potentially important component in real-life situations – the occurrence of a recall test prior to the introduction of misinformation (for exceptions, see Lane, Mather, Villa, & Morita, 2001; Saunders & MacLeod, 2002)[1]. In real-life situations, an eyewitness usually recalls the witnessed event at the crime scene, and most likely prior to encountering any misinformation. What are the implications of taking such immediate "tests" on eyewitness suggestibility?

Research suggests that taking an immediate test may reduce a person's susceptibility to misinformation. According to the traditional verbal learning literature (Erdelyi & Becker, 1974) and more recently the testing effect literature (Chan, McDermott, & Roediger, 2006), immediate recall should enhance retention of the witnessed event; thereby rendering people less susceptible to misinformation. In the present study, we conducted three experiments to test this hypothesis. In Experiment 1A, we compared the final recall performance between young adults who have completed an immediate cued recall test with those who have not taken such a test. In Experiment 1B, we attempted to replicate and extend Experiment 1A's finding to older adults, who are more prone to having erroneous recollections (Chan & McDermott, 2007a; Thomas & Sommers, 2005). In Experiment 2, we attempted to uncover the mechanisms that underlie Experiment 1A and 1B's findings.

Experiment 1A

Methods

Participants. Thirty-six undergraduate students (N = 18 for each testing condition; 20 female and 16 male) from Washington University in St. Louis participated in this experiment for course research credit.

Materials and Procedure. Subjects first viewed a ~40 min episode of the television program "24" (the witnessed event).[2] No subjects had seen this video before. Subjects in the testing condition then took an immediate recall test on 24 details of the video (e.g., Question: What did the terrorist use to knock out the flight attendant? Answer [not provided to subjects]: A hypodermic syringe), whereas subjects in the no-testing condition played Tetris (a computerized falling-rock puzzle game) for the same amount of time (12 min). During the cued recall test, subjects were told to type in their answers within 25 s of the question onset. Accuracy was encouraged and participants were not required to produce an answer. No corrective feedback was provided. Subjects in both the testing and no-testing conditions then completed a brief demographic questionnaire, a synonym and antonym vocabulary test (Salthouse, 1993), and computerized Operation Span (OSPAN) (Kane & Engle, 2003) as distractor tasks. This distractor phase lasted approximately 25 minutes.

Participants then listened to an 8 min audio narrative that described the video. They were told that the narrative was a recap of the video (the experimenter did not warn subjects about the veridicality of the narrative). Of the 24 details targeted by the immediate test, 8 of them were presented correctly in the narrative (*consistent*. E.g., [the terrorist] knocks the flight attendant unconscious with a *hypodermic syringe*), 8 were not mentioned in the narrative (*control.* E.g., [the terrorist] knocks the flight attendant unconscious), and 8 were changed in the narrative (*misleading*. E.g., [the terrorist] knocks the flight attendant unconscious with a *chloroform rag*). The misleading information always involved replacing a specific item with a plausible alternative. Each critical detail appeared only once in the narrative and whether the detail was consistent, control, or misleading was counterbalanced across subjects. Both focal and non-focal details were modified. Pilot testing ensured that all items were able to elicit false recall at a reasonable rate (M = .21). The final cued recall test was identical to the immediate test and subjects were told to report the information presented in the video. Overall, the only difference between subjects in the two testing conditions was whether the subjects have taken an immediate test.

Results and Discussion

Results are reported with an alpha level of .05 unless otherwise noted; p_{rep} indicates the probability of replicating the direction of an effect. During the immediate recall test, 63% of subjects' response were accurate recall, 3% matched the misinformation (i.e., baserate false recall), 9% were no answers or "I don't know", and 24% were responses that neither matched the correct answer nor the misinformation – these responses were classified as "Other".

More important for current purposes are the results from the final test. Surprisingly, not only did immediate testing fail to reduce the misinformation effect, it greatly increased the misinformation effect! Figure 1 displays the means of accurate and misinformation recall for subjects in the testing and no-testing conditions. The three-way interaction between testing condition (testing vs. no-testing), question type (consistent, control, misleading), and response type (correct vs. misinformation recall), was significant, $F(2, 33) = 4.51$, $p_{rep} > .93$. Subjects in the two testing conditions performed comparably on the consistent and control questions but not on the misleading questions. No ANOVA was conducted to examine the two-way interactions (between testing condition and response type) because the probabilities of correct and misinformation recall were partially complementary (see Table 1 for probabilities of responses that belong to the No answer and Other categories). A planned comparison, however, reveals that subjects who had taken the immediate test were much more likely (~20%, see the rightmost pair of bars in Figure 1) to recall the misinformation than subjects who had not taken such a test, $t(34) = 2.49$, $p_{rep} = .98$, which is *a reversed testing effect on misinformation* suggestibility!

For the control questions, accurate recall probability was at 65% in the testing condition and 62% in the no-testing condition (see the third pair of bars in Figure 1), $t < 1$, $p_{rep} = .62$. Why did the control questions fail to show a significant testing effect? Because the testing effect tends to increase with delay (Roediger & Karpicke, 2006), it is possible that the current (25 min) delay might not have been long enough for the testing advantage to be revealed, though the difference is in the expected direction.

Experiment 1B

Experiment 1B was designed to extend Experiment 1A's findings to a different subject group – older adults. Older adults are of particular interest for current purposes because they are frequently involved in eyewitness memory situations and they are particularly susceptible to the misinformation effect (Roediger & Geraci, 2007), but they are also capable of suppressing false memories with appropriate training (McDermott & Chan, 2006; Thomas & Bulevich, 2006). Therefore, it is unclear whether older adults will show the same counterintuitive effect as the younger adults.

Methods

Participants. Sixty healthy, community-dwelling older adults ($N = 30$ for each testing condition; 39 female and 21 male) participated for $15. They were recruited from the older adult subject pool maintained by the psychology department at Washington University. None of the participants have seen "24" before. The average age was 72.57 ($SD = 8.56$) for the testing group and 74.37 ($SD = 7.79$) for the no-testing group, $t < 1$. Average years of education was 15.33 ($SD = 3.46$) for the testing group and 14.00 ($SD = 2.42$) for the no-testing group, $t(58) = 1.73$, $p = .09$. Finally, average vocabulary scores was .72 ($SD = .27$) for the testing group and .58 ($SD = .27$) for the no-testing group, $t(58) = 1.94$, $p = .06$.[3]

Materials and Procedure. The materials and procedure were the same as in Experiment 1A with the following exceptions. First, subjects completed the tests on paper instead of on the computer. Second, the older adults worked on logic and lateral thinking problems (http://www.folj.com/lateral/) instead of the OSPAN task, because pilot testing revealed that the OSPAN task was too difficult for them.

Results and Discussion

Older adults recalled approximately 40% of the correct details during the initial test. Moreover, 19% of the time they did not answer a question, 7% of the time they provided an answer that matched the misinformation, and 34% of the time they provided an answer in the "Other" category.

Figure 2 displays results of the final test, which shows that older adults demonstrated the same surprising pattern as their younger counterparts. The three-way interaction between testing condition, question type, and response type was significant, $F(2, 57) = 7.86$, $p_{rep} = 99$. That is, similar to younger adults, older adults who had taken the initial test showed dramatically heightened (~20%) susceptibility to misinformation, $t(58) = 3.87$, $p_{rep} = .99$. Interestingly, they also showed a greater testing effect (on accurate recall) for the control (~9%) and consistent (~10%) questions, though neither comparisons were statistically significant, possibly due to power issues (*observed power*s < .35), both *p*s > .10, but both p_{rep}s > .79. One possible reason that older adults demonstrated a greater (though nonsignificant) testing effect than the younger adults is that they forget information at a faster rate (Floden, Stuss, & Craik, 2000), and the testing effect is typically more robust after forgetting has set in.

In sum, although older adults were much less likely to recall the correct details during the final test, they were as likely as the younger adults to recall the misinformation. Critically, both younger and older adults were more susceptible to the misinformation effect after having taken an initial recall test.

Experiment 2

Experiments 1A and 1B show that, contrary to the hypothesis we proposed based on the testing effect, initial testing can substantially enhance the misinformation effect. What potential processes could have led to such counterintuitive results? Two findings may prove helpful in further understanding this *reversed testing effect on misinformation*. The first, coming from the verbal learning literature, is termed the *insulation effect* (Tulving & Watkins, 1974). Specifically, if subjects need to learn two paired associates sequentially (e.g., A-B *and then* A-D), recall of A-D is improved if subjects have taken a test on A-B *before* learning A-D, relative to a condition in which subjects did not take a test on A-B before learning A-D. That is, prior testing enhances learning of new information. When this logic is applied to the current finding, it suggests that the reversed testing effect occurred because the testing subjects learned the misinformation better than the no-testing subjects (see also, Izawa, 1967).

Another finding that may prove helpful in explaining our results is that consolidated memories can return to a labile state after they have been reactivated (Nader, Schafe, & Le Doux, 2000; Walker, Brakefield, Hobson, & Stickgold, 2003). That is, when a consolidated memory is reactivated, it must undergo reconsolidation, during which the reactivated memory is particularly susceptible to interference (Nader et al., 2000; Walker et al., 2003). Applying this logic to the current scenario, subjects in the testing condition reactivated their memories of the witnessed event during the immediate test, thus making those memories more susceptible to interference (by misinformation) than subjects in the no-testing condition.[4]

In this experiment, we tested these possibilities in a modified-modified free recall (MMFR) design. Barnes and Underwood (1959) devised the MMFR procedure to minimize response competition when subjects need to recall multiple targets associated with a single cue (e.g., A-B, A-D). In this procedure, subjects are told to recall everything that is associated with a cue regardless of the study order. That is, when they encounter cue A, subject should recall both B and D.

When the MMFR task is used in the present context, it allows one to observe how initial testing affects learning of the misinformation. Specifically, according to the insulation effect, initial testing should enhance the recall probability of the misinformation, but it should have little effect on the recall probability of the original (correct) details. In contrast, according to the reconsolidation account, initial testing should increase the original memory's susceptibility to interference, thereby reducing the probability of accurate recall, but it should have little effect on misinformation recall.

Methods

Participants. A total of 48 undergraduate students (*N* = 24 for each testing condition; 28 female and 20 male) from Washington University participated. None of the participants have seen this particular episode of "24" within one month of the experiment, and 42 of them had never seen the first season of "24".

Procedure. All experimental protocols were the same between Experiment 2 and Experiment 1A until the final test, at which point subjects were given the MMFR instructions. Subjects were told to recall everything they could remember for each question in the final test, regardless of the accuracy or source of the memory.

Results and Discussion

For the initial test, accurate recall occupied 63% of the responses, while misinformation, no answer, and other responses occupied 5%, 5% and 27% of the responses, respectively. Figure 3 displays results of the final test (see also Table 1), which show that the reversed testing effect on misinformation can be attributed to enhanced learning of the misinformation due to initial testing (see Figure 3), which supports the insulation effect account. I now provide statistical support for this impression before turning to the reconsolidation possibility.

Similar to Experiments 1A and 1B, a 2 (testing condition) X 3 (question type) X 2 (response type) ANOVA showed a three-way interaction, $F(2, 45) = 9.12$, $p_{rep} = .99$. Of particular interest here is whether testing affected the recall probabilities of the misleading questions, it did. Specifically, there was a significant interaction between testing condition and response type, $F(1, 46) = 9.99$, $p_{rep} = .97$. A planned comparison shows that, consistent with the insulation effect prediction, subjects who had taken the initial test recalled considerably more misinformation than subjects who had not taken the initial test (21% more, see the rightmost pair of bars in Figure 3), $t(46) = 5.03$, $p_{rep} = .99$. Moreover, the magnitude of this effect closely matches the magnitude of the reversed testing effect in Experiment 1A and 1B.

No difference was found in the correct recall probabilities (for the misleading questions) between the testing subjects (.54) and the no-testing subjects (.52), $t < 1$, $p_{rep} = .30$ (see the fifth pair of bars in Figure 3). On the surface, this null effect may indicate that the reconsolidation idea fails to explain the reversed testing effect on misinformation. However, this comparison is clouded by a confound between the benefits of testing and the negative impact of interference during the reconsolidation process. That is, this null effect might have occurred because subjects in the testing condition (the gray bar) benefited from taking the initial test, but this benefit was masked by the negative impact of reconsolidation.

To examine whether our data support the reconsolidation account without this confound, we compared the level of interference experienced by subjects who had taken the initial test with those who had not. The magnitude of interference was estimated by the *within subjects* comparison between the recall probabilities of the control and misleading questions. This 2 (question type: control, misleading) X 2 (testing condition) ANOVA revealed a significant interaction between question type and testing condition, $F(1, 46) = 5.47$, $p_{rep} = .92$. Specifically, no significant interference was found for the subjects who did not reactivate their witnessed event memories before encountering the misinformation (a -3% interference effect; compare the third and fifth white bars), $t(23) = 1.07$, $p_{rep} = .65$. However, substantial interference was observed (a 9% interference effect, compare the third and fifth gray bars) for subjects in the testing condition, $t(23) = 2.16$, $p_{rep} = .89$. This pattern is consistent with the reconsolidation account, which states that recently retrieved memories become particularly susceptible to interference.

In addition to the above findings, the current experiment also showed a significant (13%) testing effect for the control details (see the third pair of bars), $t(46) = 2.68$, $p_{rep} = .95$, such that initial testing enhanced later recall if no interfering misinformation was presented. This result shows that it is possible to obtain the reversed testing effect on misinformation even when a significant, regular testing effect is found for the control questions.

General Discussion

Our results provide a first look into how immediate recall affects later eyewitness suggestibility. Contrary to the belief that immediate recall would enhance retention of details of a witnessed event and thus reduce an eyewitness' susceptibility to misinformation, here we have shown that immediate recall actually intensifies the misinformation effect (for both younger and older adults). This reversed testing effect on misinformation might be based on two mechanisms. First, initial recall enhances the learning of later misinformation and second, initial recall increases the recently recalled items' propensity to interference.

Testing Enhances Learning of Misinformation

The finding that retrieval enhances learning of subsequent misinformation joins other evidence in showing that testing has two primary effects on learning. First, testing directly enhances long-term retention of the tested materials (Chan et al., 2006; Roediger & Karpicke, 2006). Second, testing enhances learning of new materials that are presented after the initial test (Izawa, 1967; Tulving & Watkins, 1974). How can testing enhance subsequent learning of the new/misleading information? There are two possibilities. First, testing may serve to reduce the level of proactive interference that original learning asserts on new learning (Chan &

McDermott, 2007b; Szpunar, McDermott, & Roediger, 2008). One possible way for this to happen is that memory for the test episode may serve to differentiate the first encoding event (witnessed event) from the second encoding event (misinformation narrative). Second, testing may potentiate new learning by drawing attention to specific aspects of the witnessed event. When subjects encountered the misinformation narrative, the items that were presented during the initial test might "pop out" and thus were better encoded. If this were the case, then it would explain why the current pattern might be difficult to obtain if the immediate test is free recall (Lane et al., 2001), because it is possible that subjects never recalled the critical details (i.e., the details that will eventually be presented as misinformation) during the initial test.

Recently Recalled Information are More Susceptible to Interference

The present experiments have demonstrated that recently recalled information may be particularly susceptible to interference. Although this pattern is consistent with the reconsolidation idea, it does not completely rule out other potential explanations. In particular, it is possible that subjects in the testing condition did not suffer more interference *during a reconsolidation process*, but that enhanced learning of the misinformation increased the level of response competition *at retrieval* during the final test. Because initial testing enhanced learning of the misinformation, it might also increase the likelihood that recall of the misinformation would block retrieval of the original detail (Roediger, 1978), even though the MMFR procedure was supposed to minimize such blocking. Although this response competition explanation seems plausible, one piece of evidence argues against it. Specifically, if response competition at retrieval is the basis for our pattern in the MMFR procedure, then it is difficult to see how subjects in the no-testing condition showed no interference from the misinformation at all. That is, no difference was found in the recall probabilities between the items that were supposed to be interfered by misinformation (misleading details) and those that were not (control details). This finding suggests that the MMFR procedure had, to an extent, minimized the influence of response competition. Therefore, though perhaps tentative, we might have indeed obtained evidence for a reconsolidation process in human episodic memory (see also, Hupbach, Gomez, Hardt, & Nadel, 2007).

Why are memories that are undergoing reconsolidation particularly susceptible to interference? From a cognitive perspective, recently reactivated memories may enhance binding of events between the original encoding episode (the video) and the new encoding episode (the narrative) because the recently recalled memories provide a clear context for such binding to occur (Lyle & Johnson, 2006). This hypothesis, though somewhat speculative, is supported by the evidence that the hippocampus plays an important role in both reconsolidation (Rossato et al., 2007) and binding (Mitchell, Johnson, Raye, & D'Esposito, 2000).

Conclusion

Regardless of the underlying reasons for the reversed testing effect on misinformation, our results indicate that even psychologists might have underestimated the malleability of eyewitness memory. When people have recently recalled information about the witnessed event and are then exposed to misinformation, they markedly increase their susceptibility to the misinformation. These results further confirmed the notion that not only does recall indicate what we know, but it also changes what we know, and sometimes these changes can have far-reaching, and perhaps negative, consequences.

References

Barnes, J. M., & Underwood, B. J. (1959). "Fate" of first-list associations in transfer theory. *Journal of Experimental Psychology, 58*, 97-105.

Bartlett, F. C. (1932). *Remembering: A study in experimental and social psychology*: NY: Macmillan.

Chan, J. C. K., & McDermott, K. B. (2007a). The effects of frontal lobe functioning and age on veridical and false recall. *Psychonomic Bulletin & Review, 14*, 606-611.

Chan, J. C. K., & McDermott, K. B. (2007b). The testing effect in recognition memory: A dual process account. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33*, 431-437.

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.

Erdelyi, M. H., & Becker, J. (1974). Hypermnesia for pictures: Incremental memory for pictures but not words in multiple recall trials. *Cognitive Psychology, 6*, 159-171.

Floden, D., Stuss, D. T., & Craik, F. I. M. (2000). Age differences in performance on two versions of the Brown-Peterson task. *Aging, Neuropsychology, and Cognition, 2000*(7), 245-259.

Fulton, D., Kemenes, I., Andrew, R. J., & Benjamin, P. R. (2005). A single time-window for protein synthesis-dependent long-term memory formation after one-trial appetitive conditioning. *European Journal of Neuroscience, 21*, 1347-1358.

Hupbach, A., Gomez, R., Hardt, O., & Nadel, L. (2007). Reconsolidation of episodic memories: A subtle reminder triggers integration of new information. *Learning & Memory, 14*, 47-53.

Izawa, C. (1967). Function of test trials in paired-associate learning. *Journal of Experimental Psychology, 75*, 194-209.

Jacoby, L. L., Woloshyn, V., & Kelley, C. M. (1989). Becoming famous without being recognized: Unconscious influences of memory produced by dividing attention. *Journal of Experimental Psychology: General, 118*, 115-125.

Kane, M. J., & Engle, R. W. (2003). Working memory capacity and the control of attention: The contributions of goal neglect, response competition, and task set to Stroop interference. *Journal of Experimental Psychology: General, 132*, 47-70.

Keppel, G., & Underwood, B. J. (1962). Retroactive inhibition of R-S associations. *Journal of Experimental Psychology, 64*, 400-404.

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.

Lindsay, D. S., Allen, B. P., Chan, J. C. K., & Dahl, L. C. (2004). Eyewitness suggestibility and source similarity: Intrusions of details from one event into memory reports of another event. *Journal of Memory and Language, 50*, 96-111.

Loftus, E. F. (1979). The malleability of memory. *American Scientist, 67*, 312-320.

Loftus, E. F., Miller, D. G., & Burns, H. J. (1978). Semantic integration of verbal information into a visual memory. *Journal of Experimental Psychology: Human Learning and Memory, 4*, 19-31.

Lyle, K. B., & Johnson, M. K. (2006). Importing perceived features into false memories. *Memory, 14*, 197-213.

McDermott, K. B., & Chan, J. C. K. (2006). Effects of repetition on memory for pragmatic inferences. *Memory & Cognition, 34*, 1273-1284.

Mitchell, K. J., Johnson, M. K., Raye, C. L., & D'Esposito, M. (2000). fMRI evidence of age-related hippocampal dysfunction in feature binding in working memory. *Cognitive Brain Research, 10*, 197-206.

Nader, K., Schafe, G. E., & Le Doux, J. E. (2000). Fear memories require protein synthesis in the amygdala for reconsolidation after retrieval. *Nature, 406*, 722-726.

Roediger, H. L. (1978). Recall as a self-limiting process. *Memory & Cognition, 6*, 54-63.

Roediger, H. L., & Geraci, L. (2007). Aging and the misinformation effect: A neuropsychological analysis. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33*, 321-334.

Roediger, H. L., & Karpicke, J. D. (2006). The power of testing memory: Basic research and implications for educational practice. *Perspectives on Psychological Science, 1*, 181-210.

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.

Rossato, J. I., Bevilaqua, L. R. M., Myskiw, J. C., Medina, J. H., Izquierdo, I., & Cammarota, M. (2007). On the role of hippocampal protein synthesis in the consolidation and reconsolidation of object recognition memory. *Learning & Memory, 14*, 36-46.

Salthouse, T. A. (1993). Speed and knowledge as determinants of adult age differences in verbal tasks. *Journal of Gerontology: Psychological Sciences, 48*, 29-36.

Saunders, J., & MacLeod, M. D. (2002). New evidence on the suggestibility of memory: The role of retrieval-induced forgetting in misinformation effects. *Journal of Experimental Psychology: Applied, 8*, 127-142.

Szpunar, K. K., McDermott, K. B., & Roediger, H. L. (2008). Testing during study insulates against build-up of proactive interference. *Manuscript in preparation*.

Thomas, A. K., & Bulevich, J. B. (2006). Effective cue utilization reduces memory errors in older adults. *Psychology and Aging, 21*, 379-389.

Thomas, A. K., & Sommers, M. S. (2005). Attention to item-specific processing eliminates age effects in false memories. *Journal of Memory and Language, 52*, 71-86.

Tulving, E., & Watkins, M. J. (1974). On negative transfer: Effects of testing one list on the recall of another. *Journal of Verbal Learning and Verbal Behavior, 13*, 181-193.

Walker, M. P., Brakefield, T., Hobson, J. A., & Stickgold, R. (2003). Dissociable stages of human memory consolidation and reconsolidation. *Nature, 425*, 616-620.

Author Notes

We thank Katie Price for her help with participant testing. We also thank Mark McDaniel at Washington University for his unrelenting support of this project. This work was in part supported by a grant from the National Institute on Aging (#52408) to Mark McDaniel.

Footnotes

[1]In both of these studies subjects have performed retrieval practice on the witnessed event before they encountered misinformation. However, unlike the current study, Lane and colleagues' (2001) used free recall instead of a cued recall as the initial test. In Saunders and MacLeod's (2002) experiments, all subjects had taken an initial free recall test prior to the introduction of misinformation (i.e., they did not have a no-testing group).

[2]We used the first episode of the first season of "24" as the witnessed event material. The audio narrative was created by modifying the episode guide provided by Fox on www.fox.com/24.

[3]To ensure that subject selection artifacts did not contribute to the findings of Experiment 1B, we equated the older adults on vocabulary scores by selecting the 25 highest scoring individuals in the no-testing group (M = .66, SD = .21) and the 25 lowest scoring individuals in the testing group (M = .66, SD = .27), $t < 1$. Consistent with the results reported in the main text, these subjects still demonstrated a reversed testing effect on misinformation recall (M = .32 for no-testing and M = .50 for testing, t = 3.06) and a regular, but nonsignificant, testing effect on accurate recall for the control (M =.32 for no-testing and M =.37 for testing, $t < 1$) and consistent questions (M =.61 for no-testing and M =.66 for testing, $t < 1$). Similarly, we equated education level by comparing the 25 individuals with the most education in the no-testing condition (M = 14.56, SD = 2.14) to the 25 individuals with the least education in the testing condition (M = 14.28, SD = 2.70), $t < 1$. Again, there was a reversed testing effect on misinformation recall (M = .32 for no-testing and M = .51 for testing, t = 3.41) and a regular, but nonsignificant, testing effect on accurate recall for the control (M =.32 for no-testing and M =.35 for testing, $t < 1$) and consistent questions (M =.61 for no-testing and M =.66 for testing, $t < 1$). These results suggest that subject selection artifacts did not contribute to the reversed testing effect on misinformation.

[4]One assumption of this hypothesis is that memories of the video event has been consolidated to a larger extent in the no-testing condition than in the testing condition when subject encountered the interfering audio narrative -- a delay of approximately 35 minutes. Although memory consolidation is typically thought to takes hours to days (or years) to complete, recently researchers have suggested that consolidation can occur within minutes of encoding (Fulton, Kemenes, Andrew, & Benjamin, 2005).

Table 1.

Probabilities of responses that belong to the "Other" and "No Answer" categories as a function of testing condition and question type for all three experiments. Standard deviations appear in parentheses.

	Experiment 1A		Experiment 1B		Experiment 2	
	Testing	No-testing	Testing	No-testing	Testing	No-testing
Consistent Questions						
Other	.11 (.10)	.11 (.07)	.19 (.19)	.23 (.19)	.18 (.13)	.16 (.11)
No Answer	.02 (.05)	.00 (.00)	.08 (.12)	.13 (.22)	.02 (.04)	.03 (.06)
Control Questions						
Other	.26 (.17)	.30 (.18)	.35 (.22)	.41 (.20)	.25 (.14)	.34 (.17)
No Answer	.07 (.11)	.03 (.05)	.20 (.21)	.23 (.19)	.07 (.09)	.14 (.14)
Misleading Questions						
Other	.09 (.11)	.17 (.19)	.20 (.15)	.30 (.14)	.18 (.13)	.24 (.15)
No Answer	.03 (.05)	.01 (.03)	.08 (.14)	.13 (.17)	.04 (.11)	.02 (.05)

Figure Captions

Figure 1. Younger adults were substantially more likely to recall the misinformation presented in the narrative as part of the original video/witnessed event if they had taken an immediate test prior to being exposed to the misinformation than if they had not taken such an immediate test. Error bars are within-subjects *.95 CI*.

Figure 2. Similar to the results depicted in Figure 1. Older adults were also more likely to mistakenly recall the misinformation during the final test. Moreover, they also demonstrated a stronger testing effect for the consistent and control questions. Error bars are within-subjects *.95 CI*.

Figure 3. Initial testing exacerbated the misinformation effect by enhancing learning of the misinformation (compare the rightmost gray bar to its adjacent white bar) and by suppressing recall of the original/correct information (compare the third gray bar to the fifth gray bar). Error bars are within-subjects *.95 CI*.

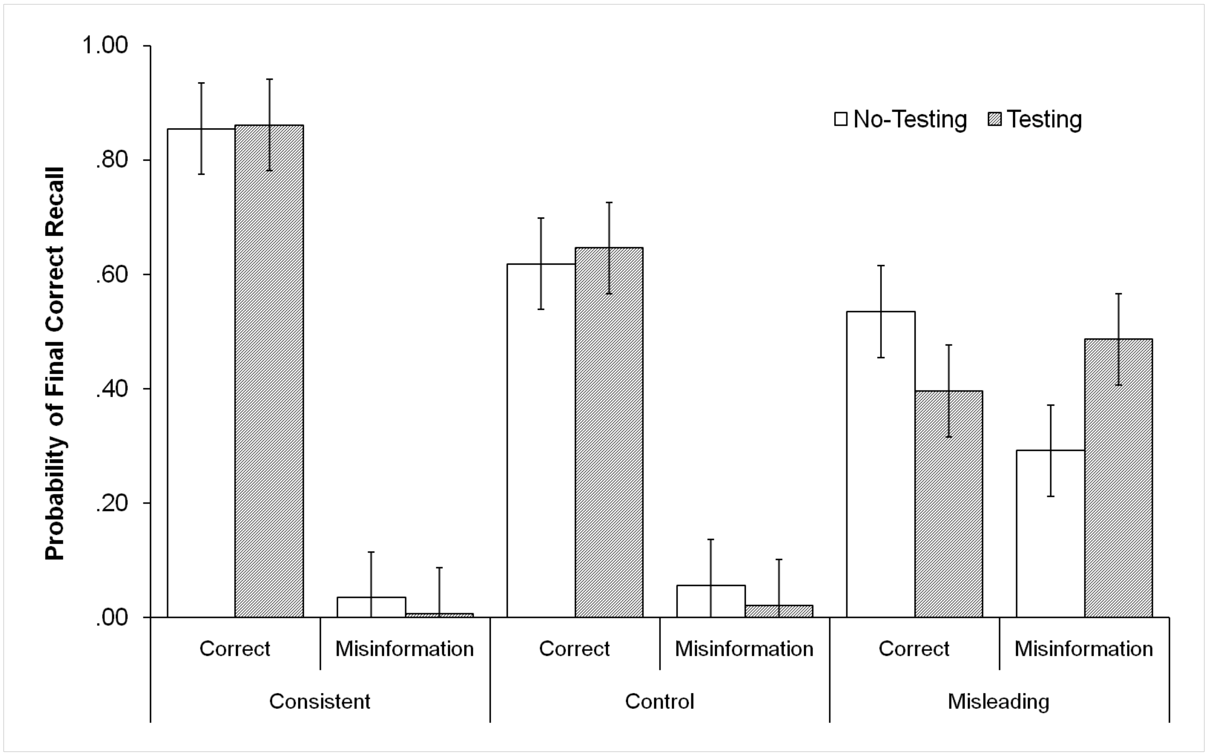
1.00
.80
.60
.40
.20
.00
Probability of Final Correct Recall
No-Testing
Testing
Correct
Misinformation
Correct
Misinformation
Correct
Misinformation
Consistent
Control
Misleading

Figure 1

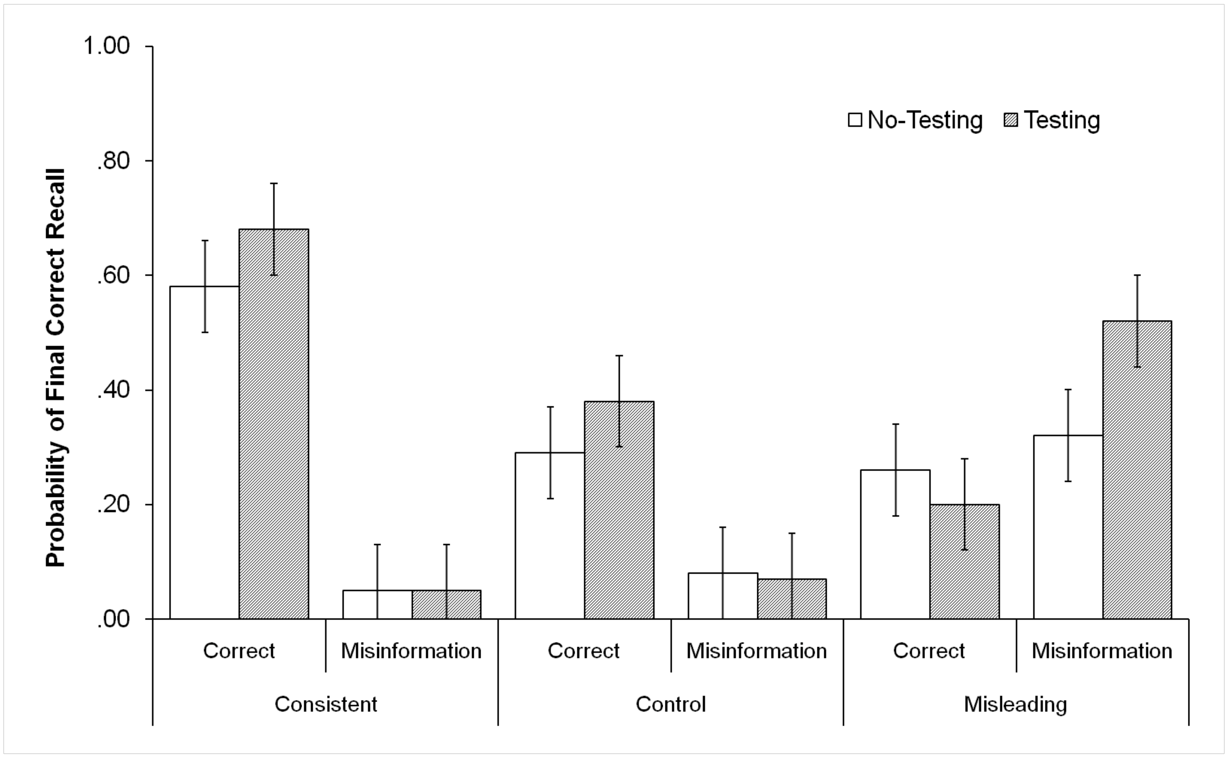
1.00
.80
.60
.40
.20
.00
Probability of Final Correct Recall
No-Testing
Testing
Correct
Misinformation
Correct
Misinformation
Correct
Misinformation
Consistent
Control
Misleading

Figure 2

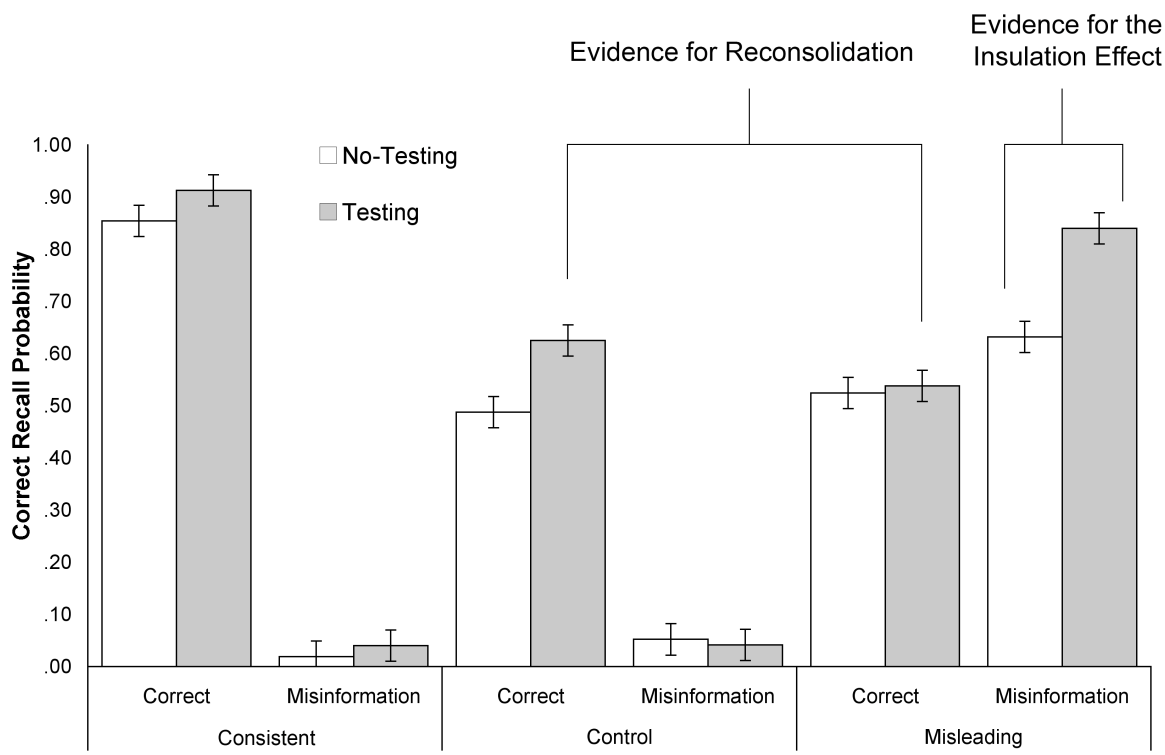
Evidence for Reconsolidation
Evidence for the Insulation Effect
1.00
.90
.80
.70
.60
.50
.40
.30
.20
.10
.00
Correct Recall Probability
No-Testing
Testing
Correct
Misinformation
Correct
Misinformation
Correct
Misinformation
Consistent
Control
Misleading

Figure 3