1-2011

Paradoxical Effects of Testing: Retrieval Enhances Both Accurate Recall and Suggestibility in Eyewitnesses

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

Moses M. Langley
University of Wisconsin - Parkside

Paradoxical Effects of Testing: Retrieval Enhances Both Accurate Recall and Suggestibility in Eyewitnesses

Abstract

Although retrieval practice typically enhances memory retention, it can also impair subsequent eyewitness memory accuracy (Chan, Thomas, & Bulevich, 2009). Specifically, participants who had taken an initial test about a witnessed event were more likely than nontested participants to recall subsequently encountered misinformation—an effect we called *retrieval-enhanced suggestibility* (RES). Here, we sought to test the generality of RES and to further elucidate its underlying mechanisms. To that end, we tested a dual mechanism account, which suggests that RES occurs because initial testing (a) enhances learning of the later misinformation by reducing proactive interference and (b) causes the reactivated memory trace to be more susceptible to later interference (i.e., a reconsolidation account). Three major findings emerged. First, RES was found after a 1-week delay, where a robust testing benefit occurred for event details that were not contradicted by later misinformation. Second, blockage of reconsolidation was unnecessary for RES to occur. Third, initial testing enhanced learning of the misinformation even when proactive interference played a minimal role. (PsycINFO Database Record (c) 2012 APA, all rights reserved)

Disciplines

Cognitive Psychology | Evidence

Paradoxical Effects of Testing: Retrieval Enhances Both Accurate Recall and Suggestibility in

Eyewitnesses

Jason C.K. Chan

Iowa State University of Science and Technology

Moses M. Langley

University of Wisconsin, Parkside

Address correspondence to:

Jason Chan
Department of Psychology
W112 Lagomarcino Hall
Iowa State University
Ames, IA 50011   USA
Email: ckchan@iastate.edu

Abstract

Although retrieval practice typically enhances memory retention, it can also impair subsequent

eyewitness memory accuracy (Chan, Thomas, & Bulevich, 2009). Specifically, participants who

had taken an initial test about a witnessed event were more likely to recall subsequently

encountered misinformation – an effect we called retrieval-enhanced suggestibility (RES). Here,

we sought to test the generality of RES and to further elucidate its underlying mechanisms. To

that end, we tested a dual mechanism account, which suggests that RES occurs because initial

testing (1) enhances learning of the later misinformation by reducing proactive interference (PI)

and (2) initial testing causes the reactivated memory trace to be more susceptible to later

interference (i.e. a reconsolidation account). Three major findings emerged. First, RES was

found after a one-week delay, where a robust testing benefit occurred for event details that were

not contradicted by later misinformation. Second, blockage of reconsolidation was unnecessary

for RES to occur. Third, initial testing enhanced learning of the misinformation even when PI

played a minimal role.

(Abstract word count: 165)

(Text word count, including References: 4989)

Paradoxical Effects of Testing: Retrieval Enhances Both Accurate Recall and Suggestibility

in Eyewitnesses

Decades of research have shown that misleading postevent information (or *misinformation*) can distort later eyewitness memory. Discovery of the misinformation effect is largely credited to Loftus and her colleagues (Loftus, Miller, & Burns, 1978), who found that subsequent memory for a previously witnessed event can be altered by exposure to intervening misleading information. In a typical misinformation experiment, participants are first presented with a witnessed event (e.g., via a video), are then exposed to misinformation about that event (e.g., via an audio narrative), and are then questioned about that event with a memory test. The general conclusion is that exposure to misinformation impairs subsequent eyewitness memory performance.

In a separate literature, researchers have rigorously examined the effects of retrieval on subsequent memory performance. The common finding from this literature is that recalling an event can enhance subsequent memory for that event – the testing effect. Cognitive psychologists have recently revitalized interest in this phenomenon and many studies have confirmed that testing is a powerful memory enhancer (e.g., Chan, McDermott, & Roediger, 2006; Karpicke & Roediger, 2008).

Although research on the testing effect and the misinformation effect have been conducted on separate fronts, it is not difficult to imagine the practical implications of the testing effect for eyewitness memory. Specifically, if recalling a witnessed event enhances an eyewitness's retention of that witnessed event, would it then reduce the likelihood that the eyewitness would fall prey to later misinformation? Recently, Chan, Thomas, and Bulevich (2009) investigated this very question. Surprisingly, they found that initial testing *increased*,

rather than decreased, participants' later report of misinformation. In this paper, we refer to this pattern of results as *retrieval-enhanced suggestibility* (RES).

The Applied and Theoretical Implications of RES

RES has important practical implications. From a legal perspective, the occurrence of RES implies that even psychologists might have underestimated the malleability of eyewitness memory. From an educational perspective, RES represents a case in which testing can *harm* memory performance. Although research has shown that testing can impair recall of *nontested* information (Anderson, Bjork, & Bjork, 1994; but see also Chan, 2009; Chan, et al., 2006), the RES effect represents a case in which testing can hurt subsequent memory for the *tested* information. Therefore, investigations of the RES effect may help provide a better understanding of when testing will help, and when it will hurt, later memory performance.

From a theoretical standpoint, investigation of RES may help further the understanding of the influence of prior retrieval on subsequent memory. Chan et al. (2009) proposed that two mechanisms are responsible for RES. The first mechanism produces RES by enhancing learning of the misinformation. Research has shown that taking a memory test on a list of paired associates (e.g., A-B) before learning a new list of paired associates (e.g., A-C) augments the learning of the latter paired associates (Tulving & Watkins, 1974), relative to a situation in which no intervening test occurs between the two learning episodes. The second mechanism leads to RES by reducing the accessibility of the witnessed event information, which occurs because testing can increase the detrimental effects of interference on a recently-retrieved memory. This idea is based on the reconsolidation notion in neurobiology (for reviews, see Dudai, 2004; Hardt, Einarsson, & Nader, 2010), which states that a memory needs to restabilize after its retrieval, and

that the memory becomes particularly susceptible to interference during this reconsolidation stage.

Placing the reconsolidation notion within the context of RES, the presentation of misinformation following initial testing might disrupt reconsolidation of the witnessed event memories, thus reducing the later recall of these original event details. In the current paper, we do not distinguish between an updating hypothesis (i.e., new learning overwrites the original memory) and an inhibition hypothesis (i.e., new learning suppresses retrieval of the original trace without overwriting it). Instead, we focus on the broader conceptualization of reconsolidation; that is, a recently-retrieved memory is particularly susceptible to interference. More generally, retrieval may enhance the assimilation of new information with an older memory (e.g., Brown, Brown, Mosbacher, & Dryden, 2006; Henkel & Franklin, 1998; Hupbach, Gomez, Hardt, & Nadel, 2007; Zaragoza & Lane, 1994).

This dual mechanism account highlights two properties of retrieval. Firstly, retrieval can enhance subsequent learning. This test-enhanced learning phenomenon is well documented in the verbal learning literature and has received renewed attention lately (Karpicke, 2009; Szpunar, McDermott, & Roediger, 2008). Secondly, encountering interference shortly after retrieval can impair retention of the retrieved item. Therefore, from a broader theoretical perspective, the RES phenomenon provides an opportunity to further understand these two unique characteristics of retrieval.

A Caveat about the RES Effect

Although a sizeable RES effect was found in Chan et al.'s experiments (2009), a regular, beneficial testing effect was absent for items that were not contradicted by later misinformation (i.e., the control items). For example, after viewing (1) a terrorist use a *syringe* to knock out a

flight attendant and (2) a police officer drive an SUV to work, a participant might hear that a

terrorist used a *chloroform rag* to knock out the flight attendant, and that a police officer drove to

work – with no misinformation being introduced for this latter detail (i.e., this was a control

item). Based on the testing effect literature, one would predict that testing should enhance

retention of at least this latter fact because it was never contradicted by misinformation.

However, Chan et al. failed to find the typical testing effect for these control items except in a

modified-modified free recall test (MMFR, Barnes & Underwood, 1959), where participants

were allowed to report multiple responses for a given query.[1]

The fact that no significant testing effect was found for the control items in Chan et al.

(2009) clouds the interpretation of RES and calls its generality into question. Indeed, the

original hypothesis – that initial retrieval might inoculate an eyewitness from future

misinformation – was predicated on the notion that testing can strengthen memory of the original

event. If testing does not enhance eyewitness memory, then there is no reason to expect that it

would reduce the misinformation effect. It is therefore crucial to investigate whether the RES

effect will occur when the benefits of initial testing on eyewitness memory is observed. The

present study therefore seeks to define the generality of RES by exploiting a characteristic of the

testing effect – namely, that it increases with delay (Chan, 2010; Wheeler, Ewers, & Buonanno,

2003). Therefore, for present purposes, investigating the RES phenomenon under a longer

retention interval might provide insight into whether RES would occur when a regular testing

effect is observed for the control questions.

The Current Experiments

In Experiments 1 and 2, participants first viewed the witnessed event, completed an

initial test or a distractor task, listened to the misinformation narrative, and then completed the

final test. In Experiment 1 (E1), the misinformation was presented immediately after the initial

test, and a 7-day retention interval separated presentation of the misinformation and the final test

(i.e., immediate misinformation). In Experiment 2 (E2), a 7-day retention interval separated the

initial test and presentation of the misinformation (i.e., delayed misinformation, and the

misinformation was presented immediately before the final test). Figure 1 displays this schedule

graphically. These experiments were designed to address one empirical question and two

theoretical questions. The empirical question, as stated earlier, concerns whether the RES effect

can co-occur with a regular testing effect for the control items. The theoretical questions concern

the dual-mechanism hypothesis, which suggests that testing exacerbates the subsequent

misinformation effect by (1) enhancing learning of the misinformation and (2) by reducing

accessibility of the witnessed event memories due to reconsolidation blockage.

A comparison of the results in E1 and E2 may shed light on this account. Specifically, if

test-enhanced interference (i.e., the reconsolidation account) is the primary contributor to RES,

then one should observe this effect in E1 but not E2. The idea is that a reactivated memory is

particularly susceptible to interference for only a short period of time, during which the memory

needs to reconsolidate (Judge & Quartermain, 1982). In E2, the misinformation was presented

one week after initial retrieval, which should provide more than enough time for reconsolidation

to occur (Berman & Dudai, 2001; Przybyslawski, Roullet, & Sara, 1999). Therefore, based on

the reconsolidation account, one would expect no RES effect in E2.

If RES is based primarily on enhanced learning of the misinformation, then there are two

predictions. To make these predictions explicit, one needs to first answer the following question:

How exactly does testing enhance subsequent learning of the misinformation? There are two

possibilities. First, testing can enhance learning of the misinformation by reducing the influence

of proactive interference (PI) on the learning of the postevent narrative (Szpunar, et al., 2008).

Because the encoding of misinformation is akin to new learning, taking an initial test on the

witnessed event might strengthen learning of the misinformation, thereby exacerbating the

misinformation effect on final recall. Second, taking a test on some details of the witnessed

event may enhance subsequent learning of these details by drawing attention to them during the

postevent narrative, which might alter participants' attentional allocation during new learning

(for a related discussion, see Nelson, Dunlosky, Graf, & Narens, 1994; Thompson, Wenger, &

Bartling, 1978). If test-reduced PI is the primary contributor to RES, then one would predict an

RES effect in only E1. Because the effects of PI on new learning diminishes with delay (e.g.,

Underwood & Freund, 1968), its impact on the learning of misinformation should be quite small

in E2, where the misinformation was presented a full week after the witnessed event. In contrast,

if RES occurs because the initial test inadvertently draws attention to the misinformation, then it

should occur in both E1 and E2. The logic is that initial testing would draw attention to the

critical details as long as a testing effect is present (or as long as people still remember the

questions presented in the initial test). Note, though, that results from the current experiments

can rule out the first two hypotheses (i.e., the reconsolidation and the PI hypotheses), but these

experiments were not designed to provide a critical test of the last hypothesis.

Experiment 1

Method

   Participants. A total of 78 Iowa State University undergraduates participated in

exchange for course credit. There were 39 participants in each testing condition.

   Materials and Procedure. Participants viewed the 43-minute pilot episode of the Fox

Television program *"24"* on a computer monitor. Audio was presented via headphones. No

participants had previously viewed this episode. Half of the participants then took a 24-question

(at 25 s each) cued recall test (e.g., Question: *What does the terrorist use on the flight attendant?*

Answer [not provided to subjects]: hypodermic syringe) while the remaining participants played

a videogame (Tetris). The initial test and the videogame both lasted approximately 12 min. All

participants then completed a demographic questionnaire and a synonym/antonym vocabulary

test for approximately five min. Next, participants listened to an eight-min audio narrative that

summarized the video. The narrative presented (a) eight *consistent* details (e.g., the terrorist

knocks a flight attendant unconscious with a *hypodermic syringe*), (b) eight *control* details (e.g.,

the terrorist knocks a flight attendant unconscious [with no mention of the critical detail]), and

(c) eight *misleading* details (e.g., the terrorist knocks a flight attendant unconscious with a

chloroform rag). The status of the critical details was counterbalanced across participants. A

week later, participants returned to complete the Operation Span (OSPAN) task and then the

final cued recall test. The final test was identical to the initial test taken by participants in the

testing condition. Participants were told to answer the questions based on their memory of the

video. See the Appendix for the instructions presented in the experiments.

Results and Discussion

A coding system classified the recall data into four categories: *Correct*, *Misinformation*,

Other, or *No Answer*. The Correct and Misinformation categories are self-explanatory. The

"Other" category refers to responses that match neither the correct answer nor the

misinformation, whereas the "No Answer" category contains skipped and "I don't know"

responses.

All analyses were conducted with an alpha level of .05. Partial eta squared (*pes*)

indicates effect size for analyses of variance (ANOVA); Cohen's *d* indicates effect size for t-

tests. Because results from the initial test are not critical for current purposes, they are reported in Table 1 only for the sake of completeness. We concentrate our analyses on the Correct and Misinformation data *from the final test*, and the data from the Other and No Answer categories are reported in Table 2.

A 2 (testing condition: no-testing, testing) X 3 (item type: consistent, control, misleading) ANOVA showed that all main effects and interactions were significant, all Fs > 10.41, pes_s > .22. Hypothesis-driven comparisons were conducted to better understand the data. First, was there a beneficial testing effect for the items that were not contradicted by later misinformation? An examination of Figure 2 reveals that the answer is a definite "Yes." Specifically, the control items revealed a substantial (22%) testing effect, $t(76) = 5.29$, $d = 1.20$, and so did the consistent items (15%), $t(76) = 3.59$, $d = .81$. Remarkably, presentation of misinformation eliminated this powerful testing advantage, $t < 1$. In fact, these items showed an RES effect, such that the tested participants were more likely to recall the misinformation ($M = .27$) than the nontested participants ($M = .17$, see the rightmost pair of bars in Figure 2), $t(76) = 2.24$, $d = .51$. Two important conclusions emerged from this finding. First, the occurrence of RES is not limited to situations in which no regular testing effect is found for control items. Second, the RES effect is long-lasting and robust.

To further examine the influence of initial testing on eyewitness suggestibility, we examined whether participants would fall prey to subsequent misinformation *if they were able to answer a question correctly during the initial test*. That is, did participants ever switch from a correct response to the misinformation from one test to the next? To our knowledge, this question is rarely answered because very few eyewitness memory studies use a test → misinformation → retest design. The results indicate that participants changed their correct

responses to the misinformation about a quarter of the time ($M = .24$, $SD = .28$), which was numerically, but not significantly, higher than the probability of misinformation recall for the nontested participants ($M = .17$), $t(76) = 1.45$, $p = .15$, $d = .32$. These data show that being able to answer a question correctly by no means inoculated participants from the influence of misinformation.

Experiment 2

Method

    Participants. A total of 60 participants (30 in each condition) participated.

    Materials and procedure. The only difference between E1 and E2 was that the 1-week delay preceded presentation of the misinformation in E2 (see Figure 1).

Results and Discussion

    The main effects for testing condition and item type and their interaction were all significant, all Fs > 3.47, pes_s > .11. Initial testing dramatically improved accurate recall for the consistent (18% better) and control (21% better) items, both ts > 3.70, ds > .95 (see the first and third pair of bars in Figure 3). However, this substantial testing benefit was again diminished for the misleading items, such that initial testing no longer significantly enhanced accurate recall for these items (see the fifth pair of bars), t < 1.07. More importantly, these items showed a powerful RES effect, such that participants who had taken the initial test reported the misinformation much more frequently ($M = .55$) than their nontested counterparts ($M = .42$), $t(58) = 2.32$, $d = .60$ (see the rightmost pair of bars). This RES effect is problematic for both the reconsolidation account and the test-reduced PI account. We explore the implications of this finding in more detail in the General Discussion.

We again examined the likelihood that participants would switch from an originally correct response to misinformation. For E2, this probability was .52 (*SD*= .35), about doubled that from E1. Again, this probability was numerically, but not significantly, greater than the misinformation recall probability of the nontested participants (*M* = .42), $t(58) = 1.29$, $p = .20$, $d = .35$. To acquire more statistical power, we combined the data from E1 and E2. This analysis revealed that the probability of switching from a correct response to the misinformation (M = .38) for the tested participants was higher than the misinformation recall probability of participants in the no-testing condition (*M* = .29), $F(1, 134) = 3.79$, *pes* = .03. This finding suggests that the RES effect is not driven only by the event details that participants failed to remember in the first place.

A comparison of E1 and E2 can also inform whether delaying presentation of the misinformation can increase its impact. Based on the literatures on verbal learning (Crowder, 1976) and on eyewitness memory (Loftus, et al., 1978; but see also Lindsay, 1990), we expected higher levels of misinformation recall in E2 than in E1 – at least in the no-testing condition. Indeed, this pattern was observed, such that the nontested participants in E2 were much more likely to report misinformation (*M* = .42) than the nontested participants in E1 (*M* = .17), $t(67) = 5.93$, $d = 1.41$. Moreover, participants in the testing condition also demonstrated greatly exacerbated misinformation recall in E2 relative to E1, $t(67) = 4.94$, $d = 1.21$ (*M* = .55 and *M* = .27 for E2 and E1, respectively). Together, these results showed that delaying the presentation of misinformation increased its influence on recall, likely because of the recency of the misinformation in E2, and this effect occurs regardless of whether an initial memory test was administered.

General Discussion

Three major findings emerged from these experiments. First, RES is a long-lasting phenomenon that can be demonstrated even when control items show a powerful testing benefit. Second, the RES effect was not confined to only items that participants were unable to recall initially. Third, blockage of reconsolidation is *not* necessary for RES to occur and testing can enhance new learning even when proactive interference plays a minimal role. We now discussed the implications of these results.

How robust is the RES effect?

A significant RES effect was found in both E1 and E2. This reveals two important properties of RES. First, the RES effect can be observed regardless of whether a significant testing effect is found for the control items. Second, the RES effect is long-lasting, such that it can be elicited a full week after participants were exposed to misinformation (E1) or a week after participants witnessed the original event (E2). Clearly, these results have important applied implications – they suggest that eyewitnesses are particularly susceptible to the misinformation effect once they have taken an immediate memory test on the witnessed event, regardless of when that misinformation is eventually encountered (at least up to a week).

Another important finding from the current experiments is that the RES effect does not appear to be driven primarily by items that participants could not recall initially. Two findings led us to this conclusion. First, even if participants could answer a question correctly on the initial test, the misinformation recall probability for these items was still greater than that of the nontested participants (when we combined the data from Experiments 1 and 2). Second, there was no difference between the rate at which participants switched from a correct ($M = .25$ in E1 and $M = .52$ in E2) and from an incorrect (including responses in the "Other" and "No Answer"

categories) response ($M = .22$ in E1 and $M = .59$ in E2) to the misinformation in both experiments, both $ts < 1$. These data are important because they highlight a powerful memory enhancing property of testing. That is, testing does not enhance learning of the misinformation for only the items that people fail to encode initially. Rather, testing can enhance learning of new information, regardless of whether that new information conflicts with an existing memory trace or whether that new information fills in a gap of the original event memory.

We elected not to issue a warning in the current experiments because eyewitnesses are virtually never warned about the potential misleading nature of postevent information in real-life situations (e.g., people can be exposed to misinformation from a variety of trustworthy sources, including the news, Lindsay, Allen, Chan, & Dahl, 2004). As a result, some participants might have treated the information presented in the narrative as correct (Belli, 1989). However, an examination of the data from Chan et al. (2009), which used the same experimental protocols, indicates that misinformation acceptance cannot account for the finding of RES. If participants mistakenly treat the postevent narrative as a source of accurate information, then they should report the misinformation whenever they remember it, even if they also remember a different piece of information from the video event. Therefore, misinformation recall probabilities should be similar between a regular cued recall test and an MMFR test. However, in Chan et al., misinformation recall probability was substantially greater in the MMFR test than in the cued recall test (an increase of .34 and .35 for the no-testing and the testing condition, respectively), suggesting that participants frequently opted to report the video event information over a corresponding misinformation, even if they remember both.[2] Moreover, the parallel increase in misinformation recall (from the cued recall test to the MMFR test) for the no-testing and testing conditions indicates that even if some level of misinformation acceptance contributes to RES, it

does not interact with testing. In other words, initial testing does not increase misinformation acceptance, even if it does increase learning of the misinformation. It remains possible, though, that the RES effect occurs only when participants are not explicitly warned about potential inaccuracies in the postevent narrative. This is an issue that future research will need to address.

Therefore, like the misinformation effect itself, the RES effect could not be attributed *solely* to strategic processes such as misinformation acceptance or demand characteristics (for a review, see Ayers & Reder, 1998). For current purposes, to the extent that research on the misinformation effect is ecologically relevant, what is important is that testing increases the misinformation effect. We now examine the theoretical underpinnings of retrieval-enhanced suggestibility.

The Dual Mechanism Account

For the dual mechanism account, the current results helped ruled out two potential mechanisms – namely, blockage of reconsolidation and test-reduced PI – as the cause of RES. As mentioned in the Introduction, both of these explanations predict that no RES should be found in Experiment 2, in which a long delay divided the initial test and the misinformation. From the perspective of the reconsolidation hypothesis, initial testing should not enhance subsequent eyewitness suggestibility if the misinformation was introduced long after the reactivated witnessed event memories had been reconsolidated. Note that although the current experiments showed that reconsolidation is not necessary for RES to occur, it remains possible that reconsolidation may augment the RES effect under some circumstances (such as those in E1). From the perspective of the test-reduced PI hypothesis, the influence of PI should have long dissipated by the time misinformation was presented in E2 (Underwood & Freund, 1968),

making it unlikely that taking an initial test provided any "advantage" on the learning of the misinformation.

The RES effect in E2 therefore leaves us with the sole remaining hypothesis. That is, initial testing enhances learning of the misinformation by allowing participants to more effectively allocate their attentional resources during the "misinformation learning phase." That is, asking participants specific questions about the complex witnessed event during the initial cued recall test might have inadvertently drawn participants' attention to these critical, and sometimes misleading, details when they listened to the audio narrative (for a related phenomenon, see Fazio & Marsh, 2008). Because the current experiments were not designed to critically evaluate this attentional allocation hypothesis, more research needs to be conducted to further examine its contribution to RES.

From an educational perspective, however, the current finding highlights the fact that prior testing can potentiate new learning (Izawa, 1970). Although previous studies have established the beneficial effects of testing on subsequent learning (e.g., Karpicke, 2009; Szpunar, et al., 2008), to our knowledge, all of them presented the new learning phase immediately or soon after initial testing.[3] Therefore, it has been unclear whether testing can potentiate subsequent learning if the two learning episodes are separated by a long retention interval – our results show that it can.

Concluding Remarks

In the current experiments, we demonstrated that initial testing can dramatically improve the long-term retention of a witnessed event. However, our data also showed that, despite a powerful, beneficial testing effect on subsequent eyewitness memory for control information, later presentation of misleading suggestions can eliminate and even reverse any such benefit.

These results are beginning to elucidate the mechanisms underlying the RES phenomenon. More generally speaking, achieving a more thorough understanding of the effects of retrieval on eyewitness memory is instrumental to helping researchers and policy makers devise methods to improve the probative value of eyewitness reports.

References

Anderson, M. C., Bjork, R. A., & Bjork, E. L. (1994). Remembering can cause forgetting: Retrieval dynamics in long-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*(5), 1063-1087.

Ayers, M. S., & Reder, L. M. (1998). A theoretical view of the misinformation effect: Predictions from an activation-based memory-model. *Psychonomic Bulletin & Review, 5*, 1-21.

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

Belli, R. F. (1989). Influences of misleading postevent information: Misinformation interference and acceptance. *Journal of Experimental Psychology: General, 118*, 72-85.

Berman, D. E., & Dudai, Y. (2001). Memory extinction, learning anew, and learning the new: Dissociations in the molecular machinery of learning in cortex. *Science, 291*, 2417-2419.

Brown, A. S., Brown, C. M., Mosbacher, J. L., & Dryden, W. E. (2006). Proactive and retroactive effects of negative suggestion. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32*, 1234-1243.

Butler, A. C., Karpicke, J. D., & Roediger, H. L. (2007). The effect of type and timing of feedback on learning from multiple-choice tests. *Journal of Experimental Psychology: Applied, 13*, 273-281.

Chan, J. C. K. (2009). When does retrieval induce forgetting and when does it induce facilitation? Implications for retrieval inhibition, testing effect, and text processing. *Journal of Memory and Language, 61*, 153-170.

Chan, J. C. K. (2010). Long-term effects of testing on the recall of nontested materials. *Memory, in press*.

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.

Crowder, R. G. (1976). *Principles of learning and memory*. Hillsdale, NJ: Erlbaum.

Dudai, Y. (2004). The neurobiology of consolidations, or, how stable is the engram? *Annual Review of Psychology, 55*, 51-86.

Fazio, L. K., & Marsh, E. J. (2008). Slowing presentation speed increases illusions of knowledge. *Psychonomic Bulletin & Review, 15*, 180-185.

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

Henkel, L. A., & Franklin, N. (1998). Reality monitoring of physically similar and conceptually related objects. *Memory and Cognition, 26*, 659-673.

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

Judge, M. E., & Quartermain, D. (1982). Characteristics of retrograde amnesia following

      reactivation of memory in mice. *Physiology & Behavior, 28*, 585-590.

Karpicke, J. D. (2009). Metacognitive control and strategy selection: Deciding to practice

      retrieval during learning. *Journal of Experimental Psychology: General, 138*, 469-486.

Karpicke, J. D., & Roediger, H. L. (2008). The critical importance of retrieval for learning.

      Science, 319, 966-968.

Lindsay, D. S. (1990). Misleading suggestions can impair eyewitnesses' ability to remember

      event details. *Journal of Experimental Psychology: Learning, Memory, and Cognition,*

      16, 1077-1083.

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., 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.

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.

Przybyslawski, J., Roullet, P., & Sara, S. J. (1999). Attenuation of emotional and nonemotional

      memories after their reactivation: role of b adrenergic receptors. *The Journal of*

      Neuroscience, 19, 6623-6628.

Szpunar, K. K., McDermott, K. B., & Roediger, H. L. (2008). Testing during study insulates

     against the build-up of proactive interference. *Journal of Experimental Psychology:*

     Learning, Memory, and Cognition, 34, 1392-1399.

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., & 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.

Underwood, B. J., & Freund, J. S. (1968). Effect of temporal separation of two tasks on proactive

     inhibition. *Journal of Experimental Psychology, 78*, 50-54.

Wheeler, M. A., Ewers, M., & Buonanno, J. F. (2003). Different rates of forgetting following

     study versus test trials. *Memory, 11*, 571-580.

Zaragoza, M., & Lane, S. M. (1994). Source misattributions and the suggestibility of eyewitness

     memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*,

     934-945.

Footnotes

[1] The MMFR procedure departs substantially from the accuracy-driven criteria that characterize most eyewitness memory reports. Consequently, these results cannot be taken as evidence that initial testing can enhance later retention of eyewitness materials under normal RES conditions.

[2] No such increase was found for any other response types. Therefore, the higher rate in misinformation recall in the MMFR test relative to the cued recall test cannot be attributed to a shift in response criterion.

[3] The delayed feedback effect has been investigated extensively (e.g., Butler, Karpicke, & Roediger, 2007), but the current issue is whether *testing* can enhance new learning after a delay. The typical studies on feedback do not include a testing vs. no-testing manipulation because feedback, by definition, is given only after a test.

Table 1

Probability of Response (Correct, Other, No Answer, or Spontaneous Recall of Misinformation)

on the Initial Recall Test. Standard deviations are presented in parentheses.

	Experiment 1 Immediate Misinformation	Experiment 2 Delayed Misinformation
Correct	.58 *(.11)*	.53 *(.10)*
Other	.31 *(.12)*	.30 *(.13)*
No Answer	.06 *(.07)*	.12 *(.11)*
Misinformation	.05 *(.04)*	.05 *(.04)*

Table 2

Probability of response ("Other" or "No Answer") as a function of testing condition and question type. Standard deviations are presented in parentheses.

| | Experiment 1 | | Experiment 2 | |
| | Immediate Misinformation | | Delayed Misinformation | |
	Testing	No-Testing	Testing	No-Testing
Consistent Questions				
Other	.24 (*.14*)	.32 (*.16*)	.12 (*.12*)	.21 (*.17*)
No Answer	.04 (*.08*)	.10 (*.14*)	.06 (*.11*)	.13 (*.16*)
Control Questions				
Other	.32 (*.16*)	.50 (*.21*)	.29 (*.19*)	.44 (*.18*)
No Answer	.06 (*.13*)	.13 (*.17*)	.15 (*.14*)	.22 (*.17*)
Misleading Questions				
Other	.27 (*.16*)	.34 (*.22*)	.19 (*.12*)	.33 (*.17*)
No Answer	.05 (*.12*)	.13 (*.14*)	.06 (*.10*)	.11 (*.14*)

Figure Captions

Figure 1. A graphical depiction of the delay schedules used in Experiments 1 and 2.

Figure 2. Results from the final test in Experiment 1: Probability of correct and misinformation recall as a function of item type (consistent vs. control vs. misleading) and initial testing condition (testing vs. no-testing).

Figure 3. Results from the final test in Experiment 2: Probability of correct and misinformation recall as a function of item type (consistent vs. control vs. misleading) and initial testing condition (testing vs. no-testing).
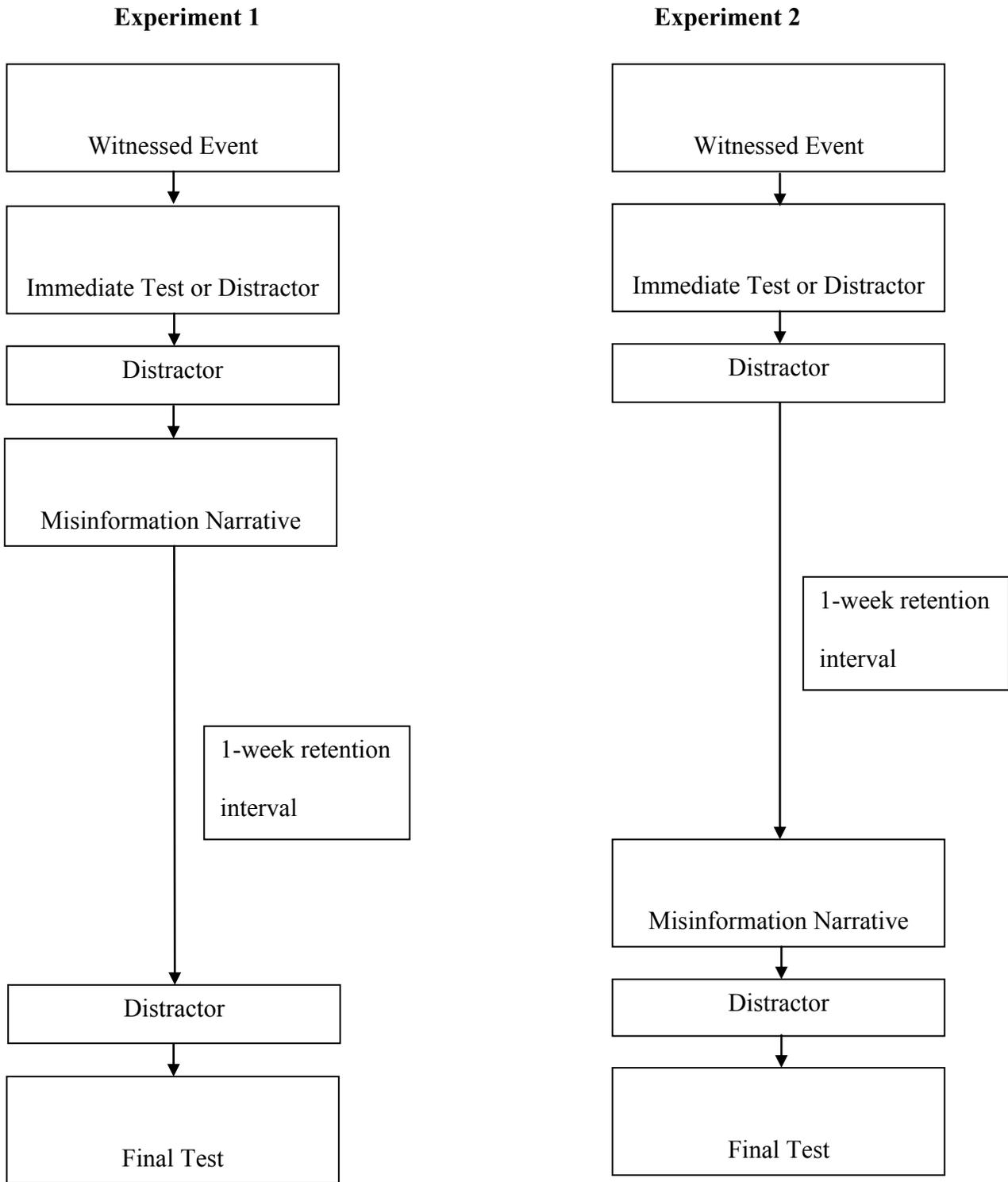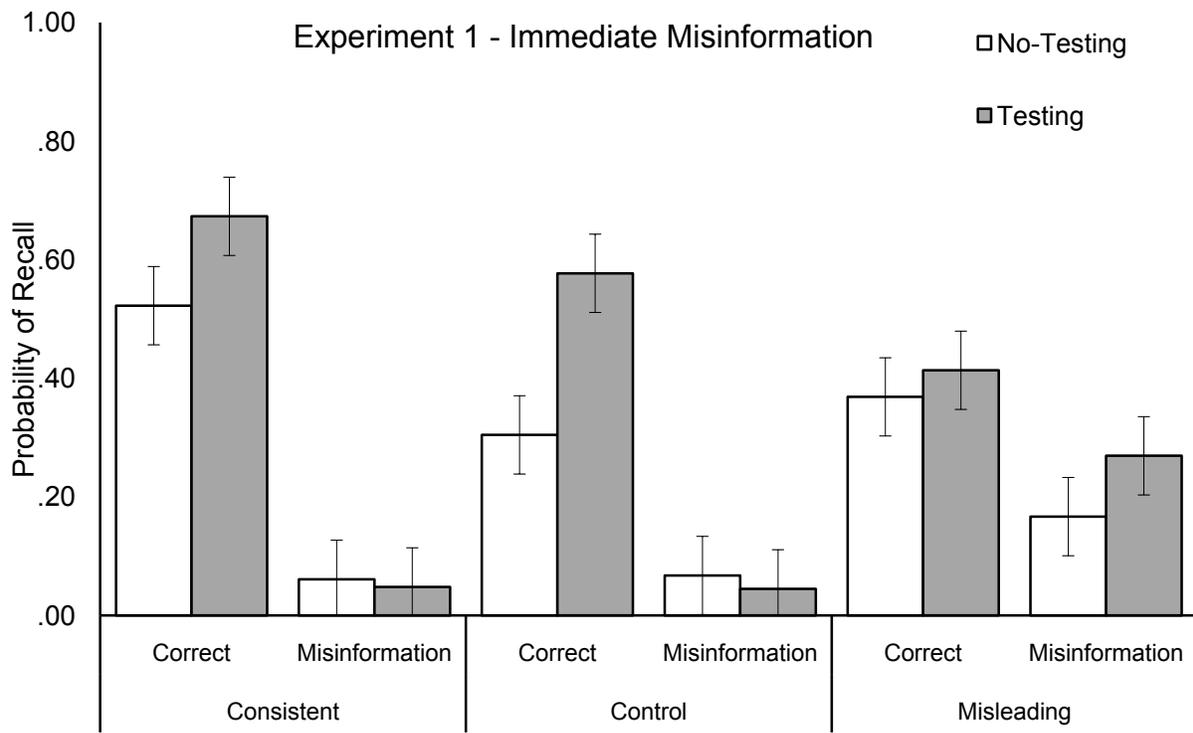
Experiment 1

```
┌─────────────────────────┐
│     Witnessed Event     │
└─────────────────────────┘
            │
            ▼
┌─────────────────────────┐
│ Immediate Test or       │
│ Distractor              │
└─────────────────────────┘
            │
            ▼
┌─────────────────────────┐
│       Distractor        │
└─────────────────────────┘
            │
            ▼
┌─────────────────────────┐
│                         │
│ Misinformation Narrative│
└─────────────────────────┘
            │           ┌──────────────────┐
            │           │ 1-week retention │
            │           │ interval         │
            │           └──────────────────┘
            ▼
┌─────────────────────────┐
│       Distractor        │
└─────────────────────────┘
            │
            ▼
┌─────────────────────────┐
│       Final Test        │
└─────────────────────────┘
```

Experiment 2

```
┌─────────────────────────┐
│     Witnessed Event     │
└─────────────────────────┘
            │
            ▼
┌─────────────────────────┐
│ Immediate Test or       │
│ Distractor              │
└─────────────────────────┘
            │
            ▼
┌─────────────────────────┐
│       Distractor        │
└─────────────────────────┘
            │           ┌──────────────────┐
            │           │ 1-week retention │
            │           │ interval         │
            │           └──────────────────┘
            ▼
┌─────────────────────────┐
│ Misinformation Narrative│
└─────────────────────────┘
            │
            ▼
┌─────────────────────────┐
│       Distractor        │
└─────────────────────────┘
            │
            ▼
┌─────────────────────────┐
│       Final Test        │
└─────────────────────────┘
```
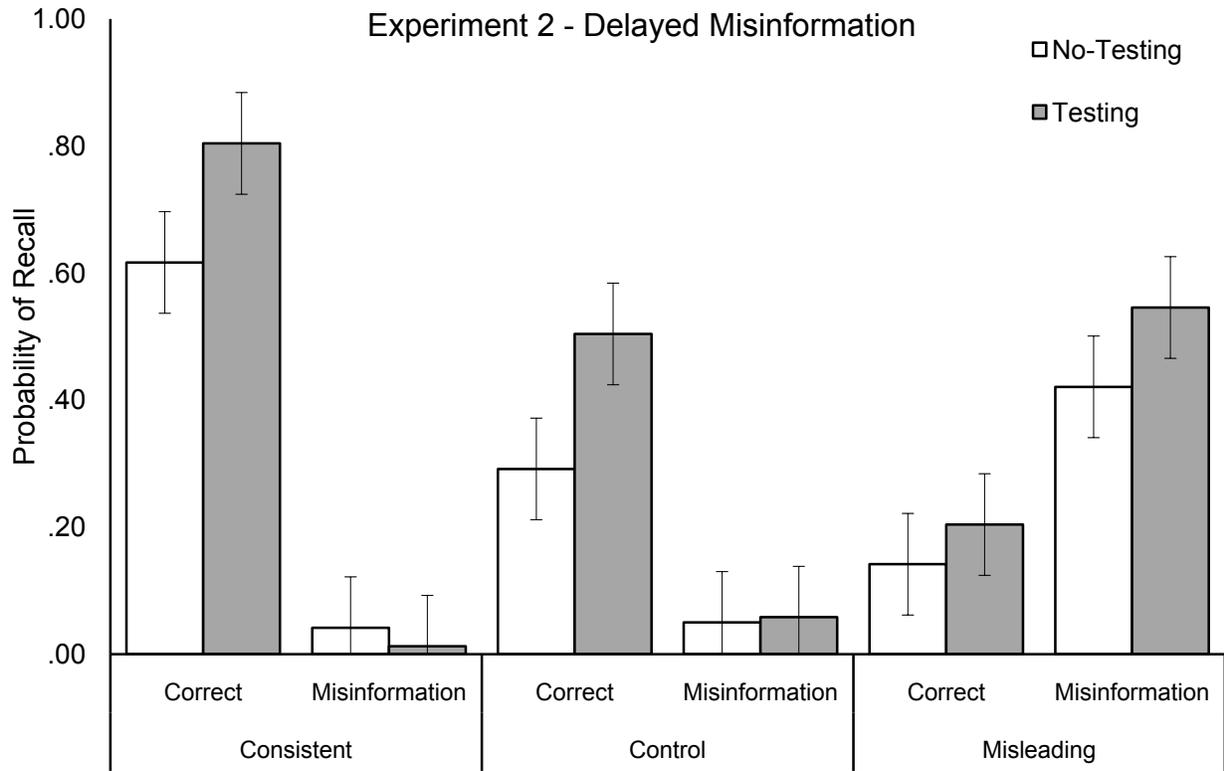
Figure 1.

Figure 2.

Figure 3.

Appendix

Instructions for the witnessed/video event

You will now watch an episode of the television show "24". The video is about 40 minutes long. Unlike watching a TV show at home, I'd like you to watch this video with a critical eye. Try to pay as much attention to the details as possible and try your best to remember what's going on in the video. Your memory for the video will be tested later.

Instructions for the initial test

Please answer the following questions to your best ability. Every question will ask you for some detail that occurred in the video you just saw. Be as accurate as possible and answer as many questions as you can. You will have 25 seconds to answer each question. There is no need to press "ENTER". The next question will show up automatically. Simply type your answer into the computer. If you have any questions, please ask your experimenter now. Otherwise, please let the experimenter know that you are ready to start.

Instructions for the audio narrative

You will now listen to an audio narrative of a recap of the episode you saw earlier. The narrative is about eight minutes long. Listen to this narrative closely. Afterwards, I'll give you a final test.

Instructions for the final test

Earlier in this experiment you watched a video. Now you will answer some questions about the video. Please answer the following questions to your best ability. Every question will ask you for some detail that occurred in the video you saw earlier. Be as accurate as possible and answer as many questions as you can. But do not guess if you have no idea what the answer is.