12-2012

Retrieval Does Not Always Enhance Suggestibility: Testing Can Improve Witness Identification Performance

Jessica A. LaPaglia
Iowa State University

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

Retrieval Does Not Always Enhance Suggestibility: Testing Can Improve Witness Identification Performance

Abstract

Verbally recalling the appearance of a perpetrator and the details of an event can sometimes hinder later eyewitness memory performance. In two experiments, we investigated the effects of verbally recalling a face on people's ability to resist subsequent misinformation about that face. Participants watched a video of a theft and then completed either a recall test or a distractor activity. After a delay, some participants heard a piece of misinformation. Memory was assessed with a recall test in Experiment 1 and with a target-present lineup in Experiment 2. In both experiments, initial testing reduced eyewitness suggestibility for the face. (PsycINFO Database Record (c) 2012 APA, all rights reserved)

Keywords

testing effect, eyewitness memory, misinformation, face recognition, verbal overshadowing, retrieval-enhanced suggestibility

Disciplines

Cognitive Psychology | Evidence

Comments

This is a manuscript of an article from Law and Human Behavior 36 (2012): 478, doi:10.1037/h0093931.

Retrieval Does Not Always Enhance Suggestibility:

Testing Can Improve Witness Identification Performance

Jessica A. LaPaglia

Jason C. K. Chan

Iowa State University

This is a manuscript of an article from *Law and Human Behavior* 36 (2012): 478, doi:10.1037/h0093931. Posted with permission. This article may not exactly replicate the final version published in the APA journal. It is not the copy of record.

Abstract

Verbally recalling the appearance of a perpetrator (Schooler & Engstler-Schooler, 1990) and the details of an event (Chan, Thomas, & Bulevich, 2009) can sometimes hinder later eyewitness memory performance. In two experiments, we investigated the effects of verbally recalling a face on people's ability to resist subsequent misinformation about that face. Participants watched a video of a theft and then completed either a recall test or a distractor activity. Following a delay, some participants heard a piece of misinformation. Memory was assessed with a recall test in Experiment 1 and with a target-present lineup in Experiment 2. In both experiments, initial testing reduced eyewitness suggestibility for the face.

Keywords: testing effect, eyewitness memory, misinformation, face recognition, verbal overshadowing, retrieval-enhanced suggestibility

Retrieval Does Not Always Enhance Suggestibility: Testing Can Improve Witness Identification Performance

When the killing of Osama bin Laden was first reported, the news was riddled with misinformation. Reports that he had been captured alive, used a woman as a human shield, and of Pakistani involvement quickly spread, despite all being false. With the growing popularity of social networking websites, the news media is competing with lay people to be the first to break news at the expense of accuracy. As a result, the scientific study of misinformation is timelier than ever.

Thanks to years of research, it is clear that eyewitness memory can be tainted by misleading postevent information (Loftus, 2005). Much of the research in this area has focused on the effects of misinformation on memory for event details (e.g., the actions performed by the perpetrator, the environment under which the incident occurred, etc.), and we have undoubtedly gained extensive knowledge in this domain. In contrast, relatively little research has investigated the effects of misinformation on memory for faces (for exceptions, see e.g., Jenkins & Davies, 1985; Loftus & Greene 1980; Searcy, Bartlett, & Memon, 2000; Windschitl, 1996). The dearth of research in this area is surprising given the vast literature on lineup identification procedure (Steblay, 1997; Wells & Bradfield, 1998) and the verbal overshadowing/facilitation effect (Meissner, Brigham, & Kelley, 2001). Generally speaking, studies in the former tradition show that witness identification accuracy can be influenced by lineup administration methods (e.g., blind vs. nonblind procedure, simultaneous vs. sequential lineup, etc., see Douglass & Steblay, 2006; Steblay, Dysart, Fulero, & Lindsay, 2001 for reviews), whereas those in the latter reveal that memory for faces can be harmed (Schooler & Engstler-Schooler, 1990) or enhanced (Brown

& Lloyd-Jones, 2005) by previous verbal recall. Clearly, similar to memory for events, memory for faces is malleable; therefore, it is important to uncover methods that can reduce eyewitness suggestibility for faces. Here, we examined whether testing, or retrieval practice, can enhance witness memory performance in a face description task and in a lineup identification task.

Eyewitness Suggestibility for Events and Faces

The current study was motivated partly by a recent counterintuitive finding (Chan, Thomas, & Bulevich, 2009), where taking a memory test about an event increased people's later susceptibility to misleading information about that event. The researchers termed this finding retrieval-enhanced suggestibility (RES). RES is surprising because initial testing was predicted to augment recollection of the original event (Pansky & Tenenboim, 2010; Roediger & Karpicke, 2006), thereby reducing suggestibility. Chan et al. suggested that RES occurs because of enhanced learning of the misinformation following retrieval. Indeed, testing has been shown to boost learning of new information (Wissman, Rawson, & Pyc, in press). Similarly, initial recall of an eyewitness event may increase the later learning of related, but misleading, information. More precisely, in Chan et al.'s experiments, participants were asked about the type of vehicle that the main character drove. Later, a narrative that contained misleading information about this detail (e.g., a pick-up truck as opposed to an SUV) was presented. When participants were asked about the vehicle during the initial memory test, this question may have inadvertently drawn their attention to the misleading detail in the narrative, thereby enhancing learning of the misinformation (similar to giving prequestions in the adjunct questions literature, Rickards & McCormick, 1988).

In the current experiments, we sought to investigate how initial retrieval affects eyewitness suggestibility for ***faces*** (as opposed to events). Research has shown that misleading postevent information can impair subsequent memory for faces. For example, Loftus and Greene (1980) found that presentation of a misleading feature (e.g., a moustache) increased the likelihood that participants would include that feature in their later description of the perpetrator. Participants were also more likely to select a foil photo with the misleading detail in a target-absent lineup (see also Searcy et al., 2000). Together, these experiments provide converging evidence for the idea that memory for faces is susceptible to the influence of misinformation.

Verbal Overshadowing and Facilitation

When no objective visual record (e.g., a closed-circuit television recording) is available, eyewitness descriptions of a perpetrator are crucial in apprehending a suspect. Research initially indicated that verbally recalling specific features of a face can hinder the subsequent accurate recognition of that face. Schooler and Engstler-Schooler (1990) termed this phenomenon verbal overshadowing. In a recent review, Chin and Schooler (2008; see also Meissner, Sporer, & Susa, 2008) identified a number of potential explanations for verbal overshadowing. For example, Meissner, Brigham, and Kelley (2001) noted that one may recall inaccurate details when asked to elaborate on a person's face. Such self-generated misinformation can impair performance on a subsequent lineup. Alternatively, verbal overshadowing can be the result of a conservative criterion shift (i.e., an increased reluctance to choose a person), because the verbal overshadowing effect disappears when participants are forced to choose a person in a target-present lineup (Clare & Lewandowsky, 2004). Lastly, verbal overshadowing may be due to recoding interference whereby verbalizing a visual memory produces a transfer inappropriate representation of the target face in memory (Dodson, Johnson, & Schooler, 1997). Notably, verbal overshadowing occurs only when the memory and test have incongruent processing demands. In particular, describing a face verbally requires participants to break down the face into individual components — such as eyes, nose, and mouth—eliciting a featural processing strategy, which differs from the way faces are normally processed (i.e., holistically). This mismatch in processing may thus harm face recognition accuracy.

Recalling a person's face, however, does not always impair later identification performance. In a meta-analysis, Meissner and Brigham (2001) found that the verbal overshadowing effect is small, though reliable (d = .12 over 29 studies). Moreover, the effect appears to be quite specific. When people are asked to *elaborate* on a face, they are more likely to show verbal overshadowing. Further, a short delay (under 30 min) between the initial test and the lineup typically results in verbal overshadowing, but delays longer that 30 min often result in verbal facilitation (which is essentially a testing effect, whereby initial testing *increases* the likelihood of correctly selecting the perpetrator later; see Schooler & Engstler-Schooler, 1990, Experiment 5, for an exception).

More recent investigations into verbal overshadowing and facilitation have uncovered some of the specific circumstances under which facilitatory effects are found. In a series of experiments conducted by Brown and Lloyd-Jones (2005), participants studied multiple faces and, after each face, described it for 15 s or completed a distractor activity. An old/new recognition task immediately followed this study/recall phase. Brown and Lloyd-Jones found verbal facilitation when participants were asked to describe similarities and differences between faces, and when they were asked to provide both holistic and featural descriptors during the initial recall phase. In a recent review, Meissner et al. (2008) explained that several factors, in addition to those described in Meissner and Brigham (2001), might contribute to verbal overshadowing and facilitation. First, verbal facilitation is more likely to occur when there are multiple targets whereas verbal overshadowing is more likely with a single target. Second, the extent to which participants are asked to provide a detailed verbal description can affect later identification accuracy. For example, when descriptions are brief and precise, verbal facilitation is more likely to occur.

The Current Study

The current study examined whether verbally recalling a face immediately following the witnessed event enhances or reduces later eyewitness suggestibility for that face. Chan et al. (2009) found that initial retrieval can enhance suggestibility for events, but it is unknown whether this effect will generalize to faces. Critically, although previous research has demonstrated the malleability of memory for both events and faces, the latter is typically considered a “special” type of stimuli (Jones, 1935), partly because of its relevance to survival and its high intraclass similarity (Werheid & Clare, 2007). Indeed, dissociable neural correlates have been identified for the encoding and retrieval of faces relative to other classes of stimuli (e.g., Chiaravallioti & Glosser, 2004; Kelley et al., 1998). It is therefore unknown whether memory for faces would be susceptible to RES.

The overall design of our experiments was modeled after those by Loftus and Greene (1980), except that we added the crucial variable of initial testing. In the first experiment, participants watched a simulated crime video and then described the perpetrator’s face (Test condition) or performed a distractor task (No-test condition). Following a 20 min delay, they listened to a narrative description of the perpetrator. The narrative included either an erroneous detail (Misleading condition) or only correct information (Control condition). After an additional 10 min delay, all participants provided a final verbal description of the suspect (i.e., the criterial test). In Experiment 2, the final recall test was replaced by a simultaneous, six-person, target-present lineup.

It is unclear whether initial testing would exacerbate or reduce the misinformation effect for face memory. One might predict that testing would lead to higher false recall based on the transfer-appropriate processing framework (Fisher & Craik, 1977; Morris, Bransford, & Franks, 1977). Specifically, because the initial recall test forces participants to transcode the target's face into a verbal description and the misinformation was also presented in a verbal description of the target, testing might therefore facilitate integration of the misinformation into the original memory. Alternatively, initial testing may reduce susceptibility to misinformation. Testing has been demonstrated to enhance delayed eyewitness identification performance (i.e., the verbal facilitation effect), so it should help participants resist the impact of misinformation. Consistent with this idea, people are sometimes less susceptible to misinformation-like manipulations when their memory for the original event is stronger (e.g., Marche, 1999; Slamecka & Ceraso, 1960; but see Bauml, 1996, Lee & Bussey, 2001, for evidence that degree of original learning has no influence on retroactive interference).

Experiment 1

Method

Participants and Design. One hundred twenty-two university undergraduate students participated in this experiment for research credit. Ten participants were excluded from analyses because English was not their primary language; therefore, all analyses were based on the remaining 112 participants (54 female, 57 male, 1 chose not to answer). The experiment used a 2 (Test vs. No-test) X 2 (postevent information: Control, Misleading) between-subjects design with 28 participants in each group.

Materials and Procedure. The experiment was conducted on computer terminals separated by dividers, with up to eight people participating simultaneously. Participants first viewed two distractor videos before the critical event video. To promote incidental encoding of the critical event video, participants were told that the videos were being evaluated as materials for a future experiment, and they needed to rate each video for its visual and audio quality. Each distractor video lasted 60 s. The first video showed a Hawaiian beach and the second showed a rabbit performing tricks.

The critical event video was approximately 45 s long and depicted a theft. A male student was shown studying in a room in the library. He was seated about 10 feet from the camera. He answers a phone call and leaves the room. A man, wearing a black shirt, then approaches. He removes a wallet from the student's backpack and steals a notebook computer before leaving. The perpetrator is a white male in his early 20s with short, brown hair and no other distinguishing characteristics. He is in view for 15 s. Following all three videos, participants were asked to rate each video on its visual and audio quality on a scale from 1 (*very poor*) to 7 (*excellent*). They then completed a demographic questionnaire. The two tasks were self-paced and lasted approximately 2-5 minutes.

Next, participants either started the initial test phase or played the video game Tetris for 10 min. Those in the Test condition were first given 5 min to type out, in as much detail as possible, a description of the target. Following this free recall test, they were given a cued recall test, which included 12 open-ended questions that targeted specific features of the perpetrator (see the Appendix for the list of questions). Participants had 25 s to answer each question. We opted for this free recall followed by specific questions test format because it resembles protocols that are used in many investigative interviews (e.g., the National Institute of Child Health and Human Development (NICHD) protocol, Orbach, Hershkowitz, Lamb, Sternberg, & Horowitz, 2000; the Cognitive Interview, Fisher & Geiselman, 1992).

Following the initial test/distractor phase, all participants completed the computerized Operation Span working memory task (Unsworth, Heitz, Schrock, & Engle, 2005), which served to prevent rehearsal of the target event and to introduce a retention interval. All participants completed this task within 30 min. Next, they listened to an audio narrative that included either one piece of misinformation (misleading narrative; that the perpetrator had facial hair on his chin) or none (control narrative). Following Loftus and Greene (1980), participants were told that a professor wrote out a description of the perpetrator immediately after watching the same video and that a research assistant recorded the description for the audio narrative. They were asked to listen to the narrative carefully without any further instructions. Participants then performed the computerized Symmetry Span working memory task (Unsworth et al., 2005), which lasted approximately 10 minutes and again served to introduce a filled retention interval. Afterwards, participants completed the final test phase, which included the same free recall followed by cued recall test as the initial test phase. The instructions emphasized to participants that their memory for the man in the video would be tested but did not mention the audio narrative. Participants were then asked, retrospectively, if they encoded the crime video intentionally (across Experiments 1 and 2, 25% of participants reported that they had intentionally memorized the critical event video).[1] Further, they were asked whether they noticed any incorrect information in the narrative (11% reported that they had noticed incorrect information in the postevent narrative, but only 5% correctly identified the misinformation).

Note that Question 8 in the cued recall test is the “critical question”. That is, we believe that participants would produce the misinformation during the final test when they attempted to answer this question. Crucially, though, that this critical question never explicitly mentioned anything about the misinformation (i.e., the facial hair). Instead, it was presented very generically: Describe any distinguishing characteristics he may have had.

Results and Discussion

The initial free recall protocol was scored based on the first 11 questions of the cued recall test. For example, if a participant reported the man's shirt color and hair color correctly but gave no other details, then that participant would have a free recall accuracy of .18 (2 out of 11 correct). For the final test, correct recall probabilities for both the free recall and cued recall tests did not include the critical question (Question 8). Therefore, correct recall probabilities for the final test were coded based on a total of 10 questions. One coder scored half of the free recall and half of the cued recall tests; a second coder scored the other half. To examine inter-rater reliability, both coders scored the final free recall data from the same 88 participants. Their coding correspondence was high, $r(86) = .95$, $p < .01$.

Analysis of variance (ANOVA) was performed to analyze continuous data (i.e., correct recall probabilities). For binary data (i.e., reporting misinformation or not), logistic regression was used to examine interaction effects, and Pearson's chi-square tests were used to examine simple and main effects. Partial eta squared (η_p^2; medium effect size ≈ .25) indicates effect size for ANOVA and Phi (φ; medium effect size ≈ .30) indicates effect size for chi-square tests.

Initial Tests. See Table 1 for results from the initial free and cued recall tests. Not surprisingly, focused questions elicited much more relevant details ($M = .61$) than free recall ($M = .27$). As expected, spontaneous reporting of the misinformation was quite rare ($M = .09$ for both tests). Thirty-nine percent of participants answered the question regarding the misinformation correctly.

Final Test. Separate analyses were conducted for the free recall and cued recall tests. See Table 2 for correct recall probabilities. A 2 (Test vs. No-test) X 2 (postevent information: Control, Misleading) ANOVA revealed no significant interaction for ***correct recall probabilities*** in either the free recall or cued recall test, *F*s < 1, *p*s > .39. As expected, no main effect of postevent information was found, *F*s < 1.60, *p*s > .20. Note that we did not anticipate the misinformation to reduce correct recall probabilities because the question regarding the misinformation (i.e., the facial hair) was not included in this set of analyses. Indeed, we did not expect the facial hair misinformation to affect recall of other aspects of the perpetrator's appearance (e.g., his hair style, how much he appeared to weigh, etc.).

There was a significant testing effect in both the free recall test, $F(1, 108) = 21.84$, $p < .01$, $\eta_p^2 = .17$, and the cued recall test, $F(1, 108) = 3.93$, $p = .05$, $\eta_p^2 = .04$. In particular, final correct recall probability was higher for participants who took the initial tests (*M* = .54 in free recall and *M* = .75 in cued recall) than for those who did not take the initial tests (*M* = .39 in free recall and *M* = .68 in cued recall). Note, however, that the testing effect in free recall might have been contaminated by the cued recall questions during the initial test phase. Because the free recall data were scored based on the cued recall questions and the tested participants had been exposed to these questions prior to the final test, these participants could therefore use the initial cued recall questions to guide retrieval during the final free recall test, giving them an additional advantage beyond just retrieval practice.

Table 3 displays the percentage of participants reporting the misinformation on the final test. A 2 (Test vs. No-test) X 2 (postevent information) logistic regression analysis revealed no interaction for both the free recall and the cued recall test, χ^2s < 1, *p*s > .56. Although no main effect of testing was found in the free recall test, $\chi^2 < 1$, $p = .68$, a significant main effect of testing was observed in cued recall, $\chi^2(1, N = 112) = 4.76, p = .03, \varphi = .21$. Specifically, testing reduced false recall from 45% to 25%. It is quite remarkable that, in the cued recall test, 75% of the participants who had been misled reported that the perpetrator had facial hair. However, when participants had taken the immediate recall test, this false recall probability dropped to a substantially lower 43%. To our knowledge, this is the first demonstration that initial retrieval can reduce eyewitness susceptibility to misleading suggestions about the facial appearance of a perpetrator.

We also examined whether answering the critical question correctly during the initial test would further inoculate participants from reporting the misinformation during the final test. As expected, participants who were correct on the initial test were less likely to report the misinformation on the final test (27%) than those who were incorrect (53%), although this difference did not reach significant, $\chi^2(1, N = 28) = 1.80, p = .18, \varphi = .25$, possibly due to low statistical power. Note that even when participants were incorrect on the initial test, they reported the misinformation at a lower rate (53%) numerically than the nontested participants (75%); again, the difference was not significant, $\chi^2(1, N = 45) = 2.32, p = .13, \varphi = .23$.[2]

In addition, there was a significant misinformation effect in both tests. In free recall, participants were far more likely to report the misinformation if they had heard it (52%) than if they had not (5%), $\chi^2(1, N = 112) = 29.58, p < .01, \varphi = .51$. A similar finding was exhibited by the cued recall test, with the misled participants more likely to report the misinformation (59%) than those who heard the control narrative (11%), $\chi^2(1, N = 112) = 28.68, p < .01, \varphi = .51$.

In sum, the results from Experiment 1 showed that initial testing can reduce the influence of misinformation on subsequent recall of a perpetrator's appearance. However, in addition to providing person descriptions to help identify a suspect, witness memory of a perpetrator is often tested in a lineup identification task. Therefore, it is important to understand whether the results from Experiment 1 would extend to such a task. We attempted to address this question in Experiment 2 using a target-present lineup.

Experiment 2

Method

Participants and Design. Two hundred fifty-seven university students participated in this experiment. Data from 17 participants were excluded from analyses: thirteen because English was not their primary language, two did not follow instructions, and two because of a computer error. Therefore, all analyses were based on the remaining 240 participants (122 females, 112 males, 6 chose not to answer). The experiment used the same 2 X 2 between-subjects design as Experiment 1, with 60 participants in each condition.

Materials and Procedure. The materials and procedure for Experiment 2 were identical to those of Experiment 1 except that participants completed a six-person, simultaneous, target-present lineup identification task as the final test. Photos included in the lineup were obtained from an Iowa State University database. Six foil photos were selected based on pilot testing to ensure relatively similar choosing rates (range: 5% to 24% on a forced-choice lineup). Specifically, the first author selected seven foil photos that matched the basic features of the perpetrator (e.g., white male in his early 20s, no facial hair, a very short haircut, and no other distinguishing features). We pilot tested these photos by having participants complete a 30-min delayed recognition task with a target-absent lineup. A computer program selected six of the seven foil photos to be included in each lineup. Based on the pilot testing results, the foil with the highest choosing rate was discarded (55% on a forced-choice lineup). All foils had no facial hair, but a version with facial hair on the chin was created digitally. In addition, a white collar was added to keep appearance consistent across all pictures, including the target picture. Five of the six foils were randomly chosen for each target-present lineup, such that each lineup contained three photos that matched the description of the misleading narrative (***Misleading Foils***) and three that matched the control narrative (one ***Target*** and two ***Other Foils***). Whether or not a foil had facial hair was randomized. The target, of course, never appeared with facial hair in the lineup.

The lineup screen contained six pictures displayed evenly in two rows. Participants were asked to look at each photo carefully and to identify the perpetrator by pressing the number that appeared beneath the photo (1-6). They could also select no photo by pressing "N," which indicated that the target was not in the lineup. The experimenter emphasized that the person in the video may or may not be present (i.e., unbiased instructions). Unbeknownst to the participants, if they selected no photo, the same lineup was displayed immediately again and they were asked to choose the person that looked most like the target. With this procedure, it is possible to examine whether non-choosers would select a misleading foil in a ***biased lineup***-like procedure.[3]

Following the lineup task, participants were asked a series of questions from Wells and Bradfield (1998), which were designed to assess the metacognitive and subjective experience associated with the lineup decision. This task was self paced. Finally, all six unaltered foil photos and the target were displayed. Participants were asked if they recognized any of the people on the screen from ***outside of the experiment*** (e.g., a friend, classmate, etc.). No participant recognized any person from the lineup.

Results and Discussion

Lineup identification was classified as Target Identification, Misleading Foil Identification (i.e., identification of a foil with facial hair), Other Foil Identification (i.e., identification of a foil without facial hair), or No Identification. Because inter-rater reliability in Experiment 1 was high, only one researcher coded the responses from the initial test in Experiment 2.

Initial Test. See Table 1 for correct and misinformation recall probabilities. Again, unsurprisingly, performance was better in cued recall than in free recall, and the baserate false recall probability was very low in both tests (*M*s = .03). Twenty-three percent of participants answered the critical question regarding the misinformation correctly.

Lineup Identification. Lineup identification data are presented in Table 4. For ***target identifications***, a 2 (Test vs. No-test) X 2 (postevent information) logistic regression analysis revealed no significant interaction, $\chi^2 < 1$, $p = .69$, but there was a significant main effect of postevent information, $\chi^2(1, N = 240) = 5.66$, $p = .02$, $\varphi = .15$, such that misinformation reduced target identifications from 23% to 12%. More important for present purposes is that there was a significant verbal facilitation (or testing) effect, $\chi^2(1, N = 240) = 5.66$, $p = .02$, $\varphi = .15$. Specifically, participants who took an initial test were more likely to select the target (23%) than those who were not initially tested (12%). This finding is consistent with verbal overshadowing studies that used at least a 30 min delay (Meissner et al., 2001). Interestingly, planned comparisons revealed that testing significantly increased correct identification when there was no misinformation, $\chi^2(1, N = 120) = 4.66$, $p = .03$, $\varphi = .20$, but this testing effect was diminished if participants had heard the misinformation, $\chi^2(1, N = 120) = 1.29$, $p = .26$, $\varphi = .10$.

For ***misinformation identifications***, there was no significant interaction between initial testing and postevent information, $\chi^2 < 1$, $p = .98$. There was, however, a significant misinformation effect, $\chi^2(1, N = 240) = 5.93$, $p = .01$, $\varphi = .16$. As expected, misinformation increased the likelihood of selecting a misleading foil (from 28% to 43%). However, there was no significant effect of initial testing on misinformation choosing. Therefore, although initial recall increased the probability of target identification in the lineup, it did not reduce the probability that one would choose a misleading foil. For ***Other foil identifications***, there was a marginally significant interaction between initial testing and postevent information, $\chi^2(1, N = 240) = 2.85$, $p = .09$. That is, testing reduced Other foil identifications in the control condition but not in the misleading condition, $\chi^2(1, N = 120) = 6.98$, $p < .01$, $\varphi = .24$.

Although witnesses are never forced to make an identification in real-life, they may nonetheless feel pressured to make an identification in a biased lineup procedure (i.e., the investigator does not inform the witness that the perpetrator may not be in the lineup). To consider the effects of initial testing and misinformation in a biased lineup, we computed the ***combined identification rates*** of the original lineup and the forced-choice lineup (see Table 5). That is, we added the selections from the forced-choice lineup (which included only nonchoosers from the original lineup) to the selections made by choosers in the original lineup.

No interaction was observed between testing and postevent information for combined ***target identifications***, $\chi^2 < 1$, $p = .93$. There was, however, a significant testing (or verbal facilitation) effect, $\chi^2(1, N = 240) = 9.09$, $p < .01$, $\varphi = .20$, with prior recall doubling target identifications (from 17% to 33%). There was also a significant main effect of postevent information, $\chi^2(1, N = 240) = 7.37$, $p < .01$, $\varphi = .18$, such that misinformation reduced target identification rate from 33% to 17%. For combined ***misinformation identifications***, there was no significant interaction between testing and postevent information, $\chi^2(1, N = 240) = 1.61, p = .20$. Planned comparisons, however, revealed that initial testing reduced misinformation choosing from 75% to 58% in the misleading condition, $\chi^2(1, N = 120) = 3.75, p = .05, \varphi = .18$. There was also a significant misinformation effect, $\chi^2(1, N = 240) = 18.21, p < .01, \varphi = .28$, such that participants were far more likely to choose a misleading foil if they had heard the misinformation (67%) than if they had not (39%).

To recap, the original lineup data suggest that when witnesses were facing a lineup with unbiased instructions, initial testing increased target identifications but it did not reduce misinformation identifications. However, when one considers the combined lineup data (which mirrors biased lineup instructions), initial testing had a dual beneficial effect on identification performance. That is, it increased target identifications and reduced misinformation identifications. The discrepancy between the original lineup data and the combined lineup data illustrates one new and very important characteristic in our results – specifically, participants who made no identification in an unbiased lineup were highly likely to pick a misleading foil – if they were in a biased situation and they did not take an initial test. Indeed, the initial nonchoosers picked a misleading foil 80% of the time during the forced-choice lineup. Remarkably, initial testing dropped this misleading foil identification rate to 47%, thus vastly reducing the misinformation effect in the combined lineup data.

Follow-up Questions. After the lineup identification task, participants were asked several follow-up questions. Because these questions were included for exploratory purposes, we did not use multiple comparison corrections when analyzing the results. Compared to nontested participants, the tested participants rated themselves as having paid less attention to the perpetrator's face, $t(238) = 2.89, p < .01, d = .37$, which is consistent with the idea that retrieval practice produces a conservative shift in metacognitive judgments (Karpicke & Blunt, 2011). However, initial testing improved participants' estimations of the duration that the target's face was visible (with testing $M = 12.99$ s, without testing $M = 10.44$ s, actual duration is 15 s), $t(238) = 1.97$, $p = .05$, $d = .26$.

In addition to examining the effects of testing on metacognitive judgments, we were interested in any potential differences in metacognitive assessments between participants who correctly identified the target and those who did not. Specifically, those who identified the target reported greater confidence and had a greater willingness to testify than those who did not (68% vs. 46% for confidence and 42% vs. 28% for willingness to testify), ts > 2.80, ps < .01, ds > .45. Furthermore, compared to incorrect participants, the correct participants more closely approximated the duration that the target's face was in view (15.59 s vs. 10.86 s), rated themselves as having a better view of the target (5.69 vs. 4.87), found the lineup task less difficult (5.26 vs. 6.30), and estimated that they spent less time on the lineup (4.55 s vs. 5.75 s), ts > 2.68, ps < .01, ds > .42. Lastly, compared to participants who did not choose the misinformation, those who did rated themselves as paying less attention to the target's face (3.80 vs. 4.30), $t(238) = 2.02$, $p = .04$, $d = .13$, and were poorer at estimating the length of time the target's face was in view (9.62 s vs. 12.78 s), $t(238) = 2.32$, $p = .02$, $d = .35$. Together, these findings are encouraging because they suggest that participants were reasonably aware of the accuracy of their choices.

General Discussion

Following a crime, witnesses are typically questioned about the details of the event and are asked to describe the perpetrator. In the present study, we found that such verbal recall may enhance overall later memory performance and protect witnesses from the harmful effects of misinformation. We now discuss our major findings in detail.

The Adverse Effects of Misleading Information on Recall and Identification

A powerful misinformation effect was found in both experiments, thus extending the findings of Loftus and Greene (1980) from a target-absent lineup to a target-present lineup. From a legal perspective, perhaps the most alarming finding here is that misinformation greatly reduced correct identifications (from 30% in the control condition to 14% in the misleading condition). Given this result, one may wonder how the effects of misinformation can be minimized. Previous research suggests that providing a warning about the credibility of the postevent source and asking people to carefully monitor the source of a retrieved item can reduce the misinformation effect for events (e.g., Lindsay, 1990; Wright, 1993; Zaragoza & Lane, 1994), although it is unclear whether these methods would be equally effective at countering misinformation for faces.

Another possible way to reduce eyewitness suggestibility for faces is to administer lineup members in a sequential manner. Wells (1993) suggested that sequential lineups can reduce the likelihood that witnesses would choose the person who looks most like the perpetrator (but see also Gronlund, 2004; Meissner, Tredoux, Parker, & Maclin, 2005, for a criterion shift account). Regardless of its underlying mechanism, sequential lineups often reduce false alarms in target-absent situations (Steblay, Dysart, Fulero, & Lindsay, 2001); it is thus possible that a sequential lineup can reduce misinformation identifications.

Although warning, deliberate source monitoring, and sequential lineups all have the ***potential*** to reduce the deleterious influence of misinformation, in the present paper, we have identified testing to be one method that can enhance eyewitness memory performance in the face of misinformation. We now discuss this finding in detail.

The Benefits of Testing on Recall and Identification

Initial testing enhanced the ability for participants to access the correct attributes of a person's face, which is manifested in the correct recall probability in both free recall and cued recall, in addition to the target identification probability in the initial lineup and combined lineup. This benefit of testing on face memory is particularly impressive when viewed in the context of recall and recognition (i.e., the identification task), which demand different contributions from controlled (i.e., recollection) and automatic retrieval processes (i.e., familiarity). Specifically, recall (especially free recall) is dominated by the recollection process whereas recognition relies more evenly on recollection and familiarity (Yonelinas, 2002). Moreover, the current recall and identification tasks differed on another dimension of their processing requirements. In particular, featural processing is essential to recalling individual components of a face, whereas holistic processing is likely instrumental to recognizing a target face among distractors (Wells & Hryciw, 1984). The fact that a testing benefit is observed across all dependent variables is a testament to its robustness. Note, however, that repeated retrieval following misinformation typically increases errors (e.g., Schreiber & Sergent, 1998).

A curious pattern in the data deserves some attention. It is not obvious why testing reduced the misinformation effect in cued recall but not in free recall in Experiment 1 (see Table 3), but a possible interpretation of this discrepancy is as follows: when the nontested participants completed the final free recall test, some of them remembered (and reported) the misinformation. When these same participants completed the following cued recall test, the specific question regarding the facial hair prompted recall of the misinformation from more participants than during the free recall task. Essentially, this interpretation is an instantiation of the well-established idea that cued recall is capable of evoking responses not available to free recall (Tulving & Pearlstone, 1966). That is, all else being equal, cued recall ***should*** lead to more misinformation production than free recall. Surprisingly, this pattern failed to materialize for the tested participants, which suggests that the initial tests helped participants resist the misinformation during the final cued recall test. Mechanistically, the final cued recall questions might have elicited participants' memory of their responses during the initial tests, which almost never contained the misinformation, and participants could therefore use this memory to resist reporting the misinformation. Note that because no specific cues were given during free recall, it would be far less likely for participants to spontaneously recall that the perpetrator had no facial hair, thus making it more difficult to reject the misinformation.

Perhaps the most important finding for current purposes is that initial testing can reduce the harmful effects of misinformation in lineup identifications. An intriguing aspect of this finding is that testing reduced the misinformation effect in the combined lineup data but not in the initial lineup data. We mentioned previously that this disparity is the result of the exceptionally high misinformation selection rate of the nonchoosers in the No-test condition, and that testing lowered these mistaken identifications in the forced-choice lineup substantially (from 80% to 47%). Impressively, testing moved a majority (68% of the 33% difference) of these would-be misinformation identifications to the target instead of the Other foils. This finding suggests that some of the nonchoosers actually recognized the target, despite their unwillingness to identify him during the initial lineup.

The effectiveness of testing in reducing the misinformation effect in a biased lineup-like procedure deserves particular attention as it has important implications for real-life eyewitness situations because biased lineup administration is thought to be a practice still used in many precincts (with Wisconsin, North Carolina, New Jersey, and several large cities being known exceptions because of their reforms in lineup administrations, www.innocenceproject.org). The fact that most of the nonchoosers in the No-test condition selected a misleading foil indicates that misinformation had its largest impact when memory of the perpetrator was weak. We discovered this result because of the implementation of the novel dual-lineup procedure.

In actual criminal cases, many factors can have a negative influence on witnesses' ability to encode or remember the appearance of a perpetrator. Indeed, many criminal activities occur quickly and under non-optimal viewing conditions. For example, the presence of a weapon (Steblay, 1992) and remote viewing distance (Loftus & Harley, 2005) can harm the ability to accurately encode the perpetrator's face. Further, many witnesses do not view a lineup until well after the event, which can seriously impair the ability to remember the perpetrator's appearance. Together, these scenarios indicate that many witnesses would be highly susceptible to the misinformation effect. Critically, testing appears to ameliorate much of the influence of misinformation, thus allowing a substantial proportion of these participants to identify the target instead of a misleading foil (23% target identification with testing and only 10% without, a greater than twofold increase, see first row of Table 5).

The Testing Effect, Verbal Facilitation, and Retrieval-Enhanced Suggestibility

The benefits of initial testing on eyewitness memory are clear from these data. Interestingly, this pattern seems to contradict those reported by Chan et al. (2009), who found that testing increased eyewitness suggestibility (i.e., RES). We suspect that major methodological differences, especially with regard to the materials, between Chan et al.'s study and ours contributed to this disparity. Chan et al.'s witnessed event lasted ~40 min whereas our critical event lasted 45 s. This is important because Chan et al. found little evidence of forgetting between the event video and the final test, so it was not possible to observe a testing effect. However, when an eyewitness event is very short, an initial test may be more effective at inoculating against forgetting and misinformation (e.g., Pansky & Tenenboim, 2010). Consistent with this idea, Chan and LaPaglia (in press) reported that conditions that produce a stronger testing effect can sometimes reduce the magnitude of RES.

Another potentially important distinction between Chan et al.'s (2009) study and ours was the instructions that accompanied the audio narrative. Whereas Chan et al.'s instruction simply mentioned that the narrative was a recap of the video event, ours mirrored those from Loftus and Greene (1980), in which participants were told that the narrative was written by a professor after he/she watched the video. This suggested that the narrative was based on the professor's memory of the video event. If participants regarded these instructions as implying that the narrative was not a verbatim representation of the video event, then they would be less susceptible to the influence of RES, because warning participants about the accuracy of the misinformation narrative can eliminate the RES effect (Thomas, Bulevich, & Chan, 2010).

Lastly, the present study investigated memory for faces whereas Chan and colleagues (2009) examined memory for events, which have different processing requirements. Faces are considered to be encoded holistically (Tanaka & Farah, 1993) whereas events are encoded sequentially (Kurby & Zacks, 2008). Due to the large differences between these materials and their underlying processing demands, it is plausible that the effects of testing and misinformation are different for these two classes of stimuli. Taking these methodological discrepancies into consideration, it is not surprising that we observed a significant testing effect here whereas Chan et al. did not.

Concluding Remarks

Attempts were made to make the current study resemble real-life eyewitness situations. For example, the event (a theft in the library) was a highly conceivable scenario for a student witness. Moreover, the format of the questioning was analogous to how an investigator might question a witness (i.e., free recall followed by specific probes). Despite our efforts, some aspects of our study may limit its generalizability. Specifically, the delay between initial retrieval and the final test was only 30 min, whereas the delay can be on the scale of weeks or months in real-life eyewitness events. Future experiments may examine how a longer delay affects lineup identification in the current paradigm. But because the benefits of testing tend to increase with retention interval (Roediger & Karpicke, 2006), it is possible that an even greater testing effect would emerge. Another possible limitation to this study is that the hit rate for nontested control participants in the unbiased lineup was quite low (15%), so it is important to examine whether the benefits of testing would vary with other viewing conditions.

In sum, the current study provided further evidence of verbal facilitation. More importantly, we uncovered a new and exciting finding given RES: initial testing can ***reduce*** suggestibility for faces. This was true in both recall and in biased-lineup identifications, though perhaps not in unbiased-lineup identifications. Therefore, the present study has identified an important boundary condition for RES.

Footnotes

[1]The data from those who reported intentionally encoding the critical event were in a similar pattern as unintentional encoders. However, because so few participants (38 and 49 participants in Experiments 1 and 2 respectively) reported intentional encoding of the event, no effects were significant.

[2] We performed the same conditional analysis for Experiment 2, but found no difference between correct and incorrect participants.

[3]In a biased lineup procedure, witnesses are not told that the perpetrator may or may not be included in the lineup. They are also not told that choosing no person is an option.

References

Bauml, K.-H. (1996). Revisiting an old issue: Retroactive interference as a function of the degree of original and interpolated learning. *Psychonomic Bulletin & Review, 3,* 380-384. doi: 10.3758/BF03210765

Brown, C. & Lloyd-Jones, T. J. (2005). Verbal facilitation of face recognition. *Memory & Cognition, 33,* 1442-1456. doi: 10.3758/BF03193377

Chan, J. C. K. & LaPaglia, J. A. (in press). The dark side of testing: Repeated retrieval can enhance eyewitness suggestibility. *Journal of Experimental Psychology: Applied.* doi: 10.1037/a0025147

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. doi: 10.1111/j.1467-9280.2008.02245.x

Chiaravallioti, N. D. & Glosser, G. (2004). Memory for faces dissociates from memory for location following anterior temporal lobectomy. *Brain and Cognition, 54,* 35-42. doi: 10.1016/S0278-2626(03)00257-4

Chin, J. M. & Schooler, J. W. (2008). Why do words hurt? Content, process, and criterion shift accounts of verbal overshadowing. *European Journal of Cognitive Psychology, 20,* 396-413. doi: 10.1080/09541440701728623

Clare, J. & Lewandowsky, S. (2004). Verbalizing facial memory: Criterion effects in verbal overshadowing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30,* 739-755. doi: 10.1037/0278-7393.30.4.739

Dodson, C. S., Johnson, M. K., & Schooler, J. W. (1997). The verbal overshadowing effect: Why descriptions impair face recognition. *Memory & Cognition, 25,* 129-139. doi: 10.3758/BF03201107

Douglass, A. B. & Steblay, N. (2006). Memory distortion in eyewitnesses: A meta-analysis of the post-identification feedback effect. *Applied Cognitive Psychology, 20,* 859-869. doi: 10.1002/acp.1237

Fisher, R. P. & Craik, F. M. (1977). Interaction between encoding and retrieval operations in cued-recall. *Journal of Experimental Psychology: Human Learning and Memory, 3,* 701-711. doi: 10.1037/0278-7393.3.6.701

Fisher, R. P. & Geiselman, R. E. (1992). *Memory-enhancing Techniques for Investigative Interviewing: The Cognitive Interview*. Springfield, IL: Thomas.

Gronlund, S. C. (2004). Sequential lineups: Shift in criterion or decision strategy? *Journal of Applied Psychology, 89,* 362- 368. doi: 10.1037/0021-9010.89.2.362

Jenkins, F. & Davies, G. (1985). Contamination of facial memory through exposure to misleading composite pictures. *Journal of Applied Psychology, 70,* 164-176. doi: 10.1037/0021-9010.70.1.164

Jones, L. W. (1935). Memory for faces and memory for names. *Journal of Personality, 4,* 115-119. doi: 10.111/j.1467-6494.1935.tb02114.x

Karpicke, J. D. & Blunt, J. R. (2011). Retrieval practice produces more learning than elaborative studying with concept mapping. *Science, 331,* 772-775. doi: 10.1126/science.1199327

Kelley, W. M., Miezin, F. M., McDermott, K. B., Buckner, R. L., Ralchie, M. E., Cohen, N. J., … Petersen, S. E. (1998). Hemispheric specialization in human dorsal frontal cortex and medial temporal lobe for verbal and nonverbal memory encoding. *Neuron, 20,* 927-936. doi: 10.1016/S0896-6273(00)80474-2

Kurby, C. A. & Zacks, J. M. (2008). Segmentation in the perception and memory of events. *Trends in Cognitive Sciences, 12,* 72-79. doi: 10.1016/j.tics.2007.11.004

Lee, K. & Bussey, K. (2001). Children's susceptibility to retroactive interference: The effects of age and degree of learning. *Journal of Experimental Child Psychology, 80,* 372-391. doi: 10.1006/jecp.2001.2638

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. doi: 10.1037/0278-7393.16.6.1077

Loftus, E. F. & Greene, E. (1980). Warning: Even memory for faces may be contagious. *Law and Human Behavior, 4,* 323, 334. doi: 10.1007/BF01040624

Loftus, E. F. (2005). Planting misinformation in the human mind: A 30-year investigation of the malleability of memory. *Learning & Memory, 12,* 361-366. doi: 10.1101/lm.94705

Loftus, G. R. & Harley, E. M. (2005). Why is it easier to identify someone close than far away? *Psychonomic Bulletin & Review, 12,* 43-65. doi: 10.3758/BF03196348

Marche, T. A. (1999). Memory strength affects reporting of misinformation. *Journal of Experimental Child Psychology, 73*, 45-71. doi: 10.1006/jecp.1998.2489

Meissner, C. A. & Brigham, J. C. (2001). A meta-analysis of the verbal overshadowing effect in face identification. *Applied Cognitive Psychology, 15,* 603-616. doi: 10.1002/acp.728

Meissner, C. A., Brigham, J. C., & Kelley, C. M. (2001). The influence of retrieval processes in verbal overshadowing. *Memory & Cognition, 29,* 176-186. doi: 10.3758/BF03195751

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

Meissner, C. A., Tredoux, C. G., Parker, J. F., & MacLin, O. H. (2005). Eyewitness decisions in simultaneous and sequential lineups: A dual-process signal detection theory analysis. *Memory & Cognition, 33,* 783-792. doi: 10.3758/BF03193074

Morris, C. D., Bransford, J. D., & Franks, J. J. (1977). Levels of processing versus transfer appropriate processing. *Journal of Verbal Learning and Verbal Behavior, 16,* 519-533. doi: 10.1016/S0022-5371(77)80016-9

Orbach, Y., Hershkowitz, I., Lamb, M. E., Sternberg, K. J., & Horowitz, D. (2000). Interviewing at the scene of the crime: Effects on children's recall of alleged abuse. *Legal and Criminological Psychology, 5,* 135-147. doi: 10.1348/135532500167930

Pansky, A. & Tenenboim, E. (2010). Inoculating against eyewitness suggestibility via verbatim vs. gist testing. *Memory and Cognition, 39,* 155-170. doi: 10.3758/s13421-010-0005-8

Rickards, J.P., & McCormick, C.B. (1988). Effect of interspersed conceptual prequestions on note-taking in listening comprehension. *Journal of Educational Psychology, 80*, 592-594. doi: 10.1037/0022-0663.80.4.592

Roediger, H. L. & Karpicke, J. D. (2006). Test-enhanced learning: Taking memory tests improves long term-retention. *Psychological Science, 17,* 249-255. doi: 10.111/j.1467-9280.2006.01693.x

Schooler, J. W. & Engstler-Schooler, T. Y. (1990). Verbal overshadowing of visual materials: Some things are better left unsaid. *Cognitive Psychology, 22,* 36-71. doi: 10.1016/0010-0285(90)90003-M

Schreiber, T. A. & Sergent, S. D. (1998). The role of commitment in producing misinformation effects in eyewitness memory. *Psychonomic Bulletin & Review, 5,* 443-448. doi: 10.3758/BF03208819

Searcy, J., Bartlett, J. C., & Memon, A. (2000). Influence of post-event narratives, line-up conditions and individual differences on false identification by young and older eyewitnesses. *Legal and Criminal Psychology, 5,* 219-235. doi: 10.1348/135532500168100

Slamecka, N. J. & Ceraso, J. (1960). Retroactive and proactive inhibition of verbal learning. *Psychological Bulletin, 57,* 449-475. doi: 10.1037/h0048601

Steblay, N. (1997). Social influence in eyewitness recall: A meta-analytic review of lineup instruction effects. *Law and Human Behavior, 21,* 283-298. doi: 10.1023/A:1024890732059

Steblay, N. M. (1992). A meta-analytic review of the weapon focus effect. *Law and Human Behavior, 16,* 413-424. doi: 0147-7307/92/0800-0413

Steblay, N., Dysart, J., Fulero, S., & Lindsay, R. C. L. (2001). Eyewitness accuracy rates in sequential and simultaneous lineup presentations: A meta-analytic comparison. *Law and Human Behavior*, 25, 459-473. doi: 0147-7307/01/1000-0459

Tanaka, J. W. & Farah, M. J. (1993). Parts and wholes in face recognition. *The Quarterly Journal of Experimental Psychology, 46,* 225-245. doi: 10.1080/14640749308401045

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. doi: 10.1016/j.jml.2010.04.004

Tulving, E. & Pearlstone, Z. (1966). Availability versus accessibility of information in memory for words. *Journal of Verbal Learning and Verbal Behavior, 5,* 381-391. doi: 10.1016/S0022-5371(66)80048-8

Unsworth, N., Heitz, R. P., Schrock, J. C., & Engle, R. W. (2005). An automated version of the operation span task. *Behavior Research Methods, 37,* 498-505. doi: 10.3758/BF03192720

Wells, G. L. (1993). What do we know about eyewitness identification? *American Psychologist, 48,* 553-571. doi: 10.1037/0003-066X.48.5.553

Wells, G. L. & Bradfield, A. L. (1998). "Good, you identified the suspect:" Feedback to eyewitnesses distorts their reports of the witnessing experience. *Journal of Applied Psychology, 83,* 360-376. doi: 10.1037/0021-9010.83.3.360

Wells, G. L. & Hryciw, B. (1984). Memory for faces: Encoding and retrieval operations. *Memory and Cognition, 12,* 338-344. doi: 10.3758/BF03198293

Werheid, K. & Clare, L. (2007). Are faces special in Alzheimer's disease? Cognitive conceptualization, neural correlates, and diagnostic relevance of impaired memory for faces and names. Cortex, 43, 898-906. doi: 10.1016/S0010-9452(08)706689-0

Windschitl, P. D. (1996). Memory for faces: Evidence of retrieval-based impairment. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22,* 1101-1122. doi: 10.1037/0278-7393.22.5.1101

Wissman, K. T., Rawson, K. A., & Pyc, M. A. (in press). The interim test effect: Testing prior material can facilitate the learning of new material. *Psychonomic Bulletin & Review*. doi: 10.3758/s13423-011-0140-7

Wright, D. B. (1993). Misinformation and warnings in eyewitness testimony: A new testing procedure to differentiate explanations. *Memory, 1*, 153-166. doi: 10.1080/09658219308258229

Yonelinas, A. P. (2002). The nature of recollection and familiarity: A review of 30 years of research. *Journal of Memory and Language, 46,* 441-517. doi: 10.1006/jmla.2002.2864

Zaragoza, M. S. & Lane, S. M. (1994). Source misattributions and the suggestibility of eyewitness memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20,* 934-945. doi: 10.1037/0278-7393.20.4.934

Table 1

Mean Probabilities of Correct and Misinformation Recall in the Initial Test Phase in Experiments 1 and 2. Standard Deviations are Presented in Parentheses

		Free Recall	Cued Recall
Experiment 1	Correct	.27 (.12)	.61 (.15)
	Misinformation	.09 (.29)	.09 (.29)
Experiment 2	Correct	.27 (.12)	.55 (.15)
	Misinformation	.03 (.18)	.03 (.16)

Table 2

Mean Probabilities of Correct Recall in the Final Test Phase as a Function of Postevent Information in Experiment 1. Standard Deviations are Presented in Parentheses

	Free Recall Final Test		Cued Recall Final Test	
	No-test	Test	No-test	Test
Control	.40 (.15)	.57 (.21)	.68 (.20)	.76 (.16)
Misleading	.38 (.14)	.50 (.15)	.68 (.14)	.74 (.20)

Table 3

Percentage of Participants Reporting the Misinformation in the Final Test Phase in Experiment 1

	Free Recall Test		Cued Recall Test	
	No-test	Test	No-test	Test
Control	7	4	14	7
Misleading	54	50	75	43

Table 4

Percentage of Participants Identifying the Target, Misleading Foils, Other Foils, and Making No Identification in Experiment 2

	Postevent Information Condition			
	Control		Misleading	
	No-test	Test	No-test	Test
Identification of Target	15	32	8	15
Identification of Misinformation	27	28	42	43
Identification of Other Foils	27	10	8	10
No Identification	32	32	42	32

Table 5

Percentage of Participants Identifying the Target, Misleading Foils, and Other Foils when the Identification Rates from the Regular and Forced-choice Lineups were Combined in Experiment 2

	Postevent Information Condition			
	Control		Misleading	
	No-test	Test	No-test	Test
Identification of Target	23	42	10	23
Identification of Misinformation	40	38	75	58
Identification of Other Foils	37	20	15	18

Appendix

Cued recall questions and the correct answers. If a question's correct answer contains more than one feature (such as that of question 10), only one was necessary to be scored as correct. Only question 8 was scored for misinformation recall.

1. What color was the man's shirt? [Correct Answer: Black]
2. What color were the man's pants? [Correct Answer: Gray]
3. Approximately how tall was he? [Correct Answer: 5'10''; response must be within 2 inches to be scored as correct]
4. Approximately how much did he weigh? [Correct Answer: 165 pounds; response must be within 10 pounds to be scored as correct]
5. What color was his hair? [Correct Answer: Brown, dirty blonde]
6. Briefly describe his hairstyle. [Correct Answer: Short, buzz-cut]
7. What shape would you say his face was? (i.e., round, oval, square, heart-shaped, etc.) [Correct Answer: Oval]
8. **Describe any distinguishing characteristics that he may have had. [Correct Answer: No distinguishing characteristics; Misinformation: Facial hair]**
9. Briefly describe his eyes (i.e., color, size, shape). [Correct Answer: Blue]
10. Briefly describe his nose (i.e., size, shape). [Correct Answer: Thin bridge, rounded tip]
11. Briefly describe his mouth (i.e., shape, thickness). [Correct Answer: Thin to average lips]
12. Describe any other details you can remember about what he looks like. [Item was not scored]