Mem Cogn (2014) 42:1239–1249
DOI 10.3758/s13421-014-0437-7

Recent study, but not retrieval, of knowledge protects against learning errors

Hillary G. Mullet · Sharda Umanath · Elizabeth J. Marsh

Published online: 28 June 2014

Abstract Surprisingly, people incorporate errors into their knowledge bases even when they have the correct knowledge stored in memory (e.g., Fazio, Barber, Rajaram, Ornstein, & Marsh, 2013). We examined whether heightening the accessibility of correct knowledge would protect people from later reproducing misleading information that they encountered in fictional stories. In Experiment 1, participants studied a series of target general knowledge questions and their correct answers either a few minutes (high accessibility of knowledge) or 1 week (low accessibility of knowledge) before exposure to misleading story references. In Experiments 2a and 2b, participants instead *retrieved* the answers to the target general knowledge questions either a few minutes or 1 week before the rest of the experiment. Reading the relevant knowledge directly before the story-reading phase protected against reproduction of the misleading story answers on a later general knowledge test, but retrieving that same correct information did not. Retrieving stored knowledge from memory might actually enhance the encoding of relevant misinformation.

Keywords Knowledge · False memory · Retrieval · Suggestibility · Fiction

Memory is surprisingly malleable. This claim is supported by many demonstrations in the episodic memory literature (i.e., memory for specific events and experiences), in which people can be easily led to misremember the details of past events and even to report entire events that never occurred (e.g., Loftus, 1975; Loftus & Pickrell, 1995). Less clear, however, is whether this same property of malleability applies to the knowledge base. General facts about the world (e.g., the Pacific is the largest ocean) have often been repeatedly encountered and may be associated with supporting knowledge in memory, helping to stabilize access to that information (Myers, O'Brien, Balota, & Toyofuku, 1984). Indeed, following a period of initial decline, access to knowledge of high school classmates' names (Bahrick, Bahrick, & Wittlinger, 1975), algebra (Bahrick & Hall, 1991), and Spanish vocabulary (Bahrick, 1984) remains remarkably stable over a period of 50 years, suggesting that, in contrast to episodic memories, stored knowledge about the world may be more resistant to change (see also Tulving, 1983).

In light of this possibility, it is striking that people learn errors that contradict their stored knowledge. For example, reading stories that contain obviously incorrect assertions (e.g., *Exercise weakens the heart and lungs* or *Seatbelts do not save lives*) leads participants to later rate these statements as truer (Prentice, Gerrig, & Bailis, 1997; Rapp, Hinze, Kohlhepp, & Ryskin, 2013; Wheeler, Green, & Brock, 1999). Similarly, readers will reproduce incorrect ideas that they have read in fictional texts (e.g., *The Atlantic is the largest ocean*), even when those inaccuracies contradict well-known facts (e.g., Marsh, Meade, & Roediger, 2003). Importantly, learners believe that they knew the misinformation before the experiment, indicating that it has been integrated with prior knowledge (Marsh et al., 2003).

If such learning of errors occurred only when people did not have the correct knowledge stored in memory, it would be unsurprising. Without prior knowledge, these would simply be examples of new episodic learning. The effects just described, however, occur even when people have prior knowledge of the correct answers, as has been shown on the basis of norms (Marsh et al., 2003) or a preexperimental (Fazio, Barber, Rajaram, Ornstein, & Marsh, 2013) or postexperimental (Bottoms et al., 2010) knowledge check. One possibility is that such effects occur because not all of

H. G. Mullet (✉) · S. Umanath · E. J. Marsh
Psychology & Neuroscience, Duke University, Box 90086, Durham, NC 27708-0086, USA
e-mail: hgm8@duke.edu

the correct knowledge that is stored, or *available*, in memory is *accessible* under every circumstance (Bahrick & Hall, 1991; Bahrick & Phelps, 1988; Tulving & Pearlstone, 1966). In other words, despite knowing the correct information that should prevent them from learning these errors, people may not be able to retrieve that information at the time that they need to apply it.

Consistent with such a possibility, Bahrick and colleagues (Bahrick & Hall, 1991; Bahrick & Phelps, 1988) coined the term *marginal knowledge* to refer to information (i.e., vocabulary and general knowledge facts) that people can correctly recognize but not recall. The idea is that items that are stored in the knowledge base but have not been recently encountered fluctuate downward in retrievability (Bahrick & Hall, 1991; Bahrick & Phelps, 1988). Importantly, a simple intervention—for example, 5 s of studying these marginal items—is enough to correct this downward fluctuation, greatly increasing the likelihood of being able to recall those items later on (Berger, Hall, & Bahrick, 1999).

Tip-of-the-tongue states, or the inability to produce a particular word or answer despite being confident that one knows the correct information, provide another demonstration that information that is stored in the knowledge base can nonetheless be inaccessible in a given moment. Importantly, tip-of-the-tongue states are less likely to occur for items that have been recently encountered (Brown, 2012; Cleary, 2006; Cleary & Reyes, 2009; Cleary & Specker, 2007; Rastle & Burke, 1996). These findings are also consistent with the claim that information that has recently been reviewed is more likely to be retrievable a short time later.

An influential idea inspired by Estes (1955a, 1955b), and later revived by Bjork and Bjork (1992) in their new theory of disuse, sheds some theoretical light on all of these findings. According to the theory, representations in memory are characterized in terms of both their *storage strength*, or how well they are learned, and their *retrieval strength*, the momentary ease of accessing those items. Both storage and retrieval strength are thought to increase with additional exposure to a given item. Once accumulated, however, storage strength is never lost, whereas retrieval strength decreases across time due to interference from competing items. From this perspective, even well-learned items can become difficult to retrieve when they have not been recently reviewed (especially when interfering items have also been encountered). Conversely, reviewing an item increases its retrieval strength, thereby increasing the likelihood that that item will be easily retrievable in the near future.

In the present research, we applied these principles to further investigate the surprising finding that people often incorporate errors into their knowledge bases, even when they have the correct knowledge stored in memory (e.g., Bottoms, Eslick, & Marsh, 2010; Fazio et al., 2013; Kelley & Lindsay, 1993; Marsh et al., 2003). We hypothesized that such learning of errors might occur because people have not recently used the relevant knowledge needed to catch the error (and therefore, the accessibility of this knowledge may be relatively low). Following from this assumption, we predicted that giving participants the opportunity to review their correct knowledge immediately before exposure to the misleading references might protect them from later reproducing that misinformation. We tested this hypothesis by requiring participants to review knowledge at a short (a few minutes) or long (1 week) delay before reading stories that contained relevant misleading information. At the short delay of just a few minutes, participants' knowledge should be highly accessible, and thus we expected that they would be protected from reproducing the misleading story answers on a later general knowledge test (relative to participants in the long-delay condition, whose knowledge should no longer be as accessible by the time of the story-reading phase). Across experiments, we also examined whether the way in which participants reviewed their knowledge had implications for later suggestibility. In Experiment 1, participants studied the target information at the beginning of the experiment, similar to the interventions typically used in the marginal knowledge and tip-of-the-tongue literatures described above. Note that this procedure also parallels that of classic eyewitness misinformation studies in the episodic memory literature, in which participants first witness an event (i.e., the correct information) and are later exposed to relevant misleading details (e.g., Loftus, Miller, & Burns, 1978). In Experiments 2a and 2b, we explored what we believed would be an even more potent method for heightening the accessibility of correct knowledge: requiring participants to *retrieve* the answers to the target questions at the beginning of the experiment.

Experiment 1

Method

Participants

Forty-eight Duke University undergraduates ($n = 26$ for the short-delay condition and 22 for the long-delay condition) participated for course credit or monetary compensation.

Design

The design was a 2 (delay: short, long) × 2 (fact frame: neutral, misleading) mixed factorial. The length of the delay between the initial study phase and the story-reading phase was the between-subjects variable, and the framing of the story references was the within-subjects variable.

Materials

For the initial study phase, participants read 64 questions (36 critical questions and 28 filler questions), along with their correct answers. These questions were selected to have a range of difficulty (1 %–85 % answered correctly on a cued-recall test) based on a recent sample of 232 Duke University undergraduates (0 %–79 % correct based on the Tauber, Dunlosky, Rawson, Rhodes, & Silzman, 2013, norms and 3 %–84 % correct based on the Nelson & Narens, 1980, norms). Sample questions included "What is the largest ocean in the world?" and "What is the last name of the woman who began the profession of nursing?"

The two fictional stories (modified from Marsh, 2004) contained characters, dialogue, and a plot and were about 1,400 words each. They were presented on the computer using MediaLab and DirectRT software (Jarvis, 2008a, 2008b). Eighteen factual references appeared in each story. These references corresponded to the 36 critical questions from the initial study phase; for each participant, half of these items were presented in a misleading frame, and half were presented in a neutral frame. Facts presented in a misleading frame suggested a plausible but incorrect answer (e.g., *paddling across the largest ocean, the Atlantic*) to the corresponding general knowledge question, whereas those presented in a neutral frame did not suggest a specific answer (e.g., *paddling across the largest ocean*). The purpose of presenting some neutral-framed facts was to provide a baseline for comparison to the misleading-framed facts.[1] In other words, the key analyses examined the tendency to produce a correct (or misleading) answer on the final general knowledge test after reading neutral versus misleading references in the stories.

The final general knowledge test took place at the end of the experiment and contained the same short-answer questions (36 critical and 28 filler) as the initial study phase.

Procedure

After giving informed consent, participants completed the initial study phase. Each general knowledge question appeared at the top of the computer screen, and the correct answer appeared at the bottom of the screen. The answer appeared at the same time as the question to reduce the possibility that participants would engage in covert retrieval. Participants were instructed to read each question and its answer and to press the Enter key when they were ready to move on. After the initial study phase, participants in the short-delay condition completed brainteasers for 3 min. Those in the long-delay condition were dismissed and reminded to return 1 week later for the second session.

After the short or long delay, participants began the story-reading phase on the computer. The instructions stated that they would be reading and listening to some fictional stories. Importantly, participants were warned that authors of fiction often take liberties with facts, so some of the story information might be incorrect. Participants read/heard several practice sentences, one of which contained a factual error. The story sentences appeared one at a time and were accompanied by a voiceover. Participants pressed the Enter key when they were ready to move on. To ensure attentiveness, 10 catch trials were included on sentences that did not contain critical factual references. On these trials, participants were prompted to type the sentence that they had just read/heard. Processing each story took about 15 min. Afterward, participants solved unrelated brainteasers for 3 min.

Participants then took the final short-answer general knowledge test on the computer. In order to emphasize that this test was a separate section of the experiment, participants received a new set of instructions after completing the puzzles. In addition, the final general knowledge test was presented using a different background color and font than the rest of the experiment. Again, participants were instructed not to guess and were allowed to answer "I don't know." After the final test, participants were debriefed, and they rated how surprising they found each of the corrected versions of the previously misleading facts. The entire experiment took about 1 h 30 min (or, for participants in the long-delay condition, about 30 min for the first session and 1 h for the second session).

Results

Data scoring

Two independent coders scored the responses to the critical questions (Cohen's kappa = .97); a third coder resolved the discrepancies. Each answer was coded as correct, misinformation, other wrong, or "I don't know." For example, for the question about the largest ocean, "Pacific" would be coded as correct, "Atlantic" as misinformation, and "Indian" as another wrong answer.

Final test performance

Our first interest was in whether participants reproduced the misleading answers from the stories. In fact, suggestibility was robust on the final test, $F(1, 46) = 57.54$, $MSE = .004$, $p < .001$, $\eta^2 = .44$; even though they had reviewed all of the correct information during the initial study phase, participants still reproduced some of the misleading references later on (M = .13 after exposure to misleading references vs. .03 after exposure to neutral references). The critical question was whether

[1] Some previous work has also included correct-framed references (e.g., Marsh & Fazio, 2006; Marsh et al., 2003). In the present research, however, our primary interest was in the reproduction of misinformation; thus, we included neutral and misleading references only.

studying the correct answers immediately before the story-reading phase, making those answers highly accessible, provided any protection against this tendency to later reproduce the story misinformation (relative to when participants had studied those answers one week before; see Fig. 1). A significant interaction between delay and fact-frame, $F(1, 46) = 27.03$, $MSE = .004$, $p < .001$, $\eta^2 = .21$, revealed that the answer to this question was yes: participants in the long-delay condition were more likely to produce the misleading answers after reading the misleading references ($M = +.17$), whereas those in the short-delay condition were only slightly more likely to do so [$M = +.03$; although the difference between the misleading and neutral conditions was still significant, even for the short-delay condition, $t(25) = 2.51$, $p = .02$, $d = 0.83$].

A second (and complementary, although not entirely independent) measure of suggestibility is whether the misinformation interfered with participants' abilities to correctly answer questions on the final test, even though they had previously reviewed all of the correct answers. In fact, participants were less likely to produce the correct answers on the final test after reading the misleading references ($M = .60$) than after reading neutral references ($M = .66$), $F(1, 46) = 11.56$, $MSE = .007$, $p = .001$, $\eta^2 = .19$. In other words, participants were no longer able to produce some of the correct information that they likely knew (and had studied in the first phase of the experiment). The interaction between delay and fact-frame was not significant, $F(1, 46) = 2.70$, $MSE = .007$, $p = .11$, $\eta^2 = .05$, although the decrease in correct responding as a result of misinformation exposure was numerically greater when participants had studied the correct answers one week earlier ($M = .09$) than when they studied those answers immediately before the story-reading phase ($M = .03$). That is, when knowledge was more accessible, participants suffered fewer decrements to their correct answers.

Note that the above analyses included both items for which participants likely knew the correct answers before the experiment (i.e., stored knowledge) and items that they did not already know (i.e., new learning). It is possible that the suggestibility effect was driven mainly by those items that represented new learning from the stories. To investigate this possibility, we separated the items into those for which participants were *more likely* to already have correct stored knowledge (the 37 % of the questions that were most often answered correctly by a Duke undergraduate sample of 232 participants; $M = .62$ correct) and those for which they were *less likely* to have correct stored knowledge ($M = .14$ correct based on the same sample).[2] We then conducted a 2 (delay: short, long) × 2 (fact-frame: neutral, misleading) × 2 (estimated stored knowledge: yes, no) ANOVA on final misinformation production. The only significant effects were the main effect of fact frame, $F(1, 46) = 55.89$, $MSE = .01$, $p < .01$, $\eta^2 = .44$, the main effect of delay, $F(1, 46) = 21.82$, $MSE = .02$, $p < .01$, $\eta^2 = .32$, and the fact-frame × delay interaction, $F(1, 46) = 25.74$, $MSE = .01$, $p < .01$, $\eta^2 = .20$. Importantly, there were no significant effects involving estimated stored knowledge (all Fs < 1); participants were just as likely to reproduce the misleading information when they probably had stored knowledge for the correct information as when they probably did not ($M = .13$, $SD = .14$, and $M = .13$, $SD = .17$, for estimated stored knowledge and no estimated stored knowledge, respectively). In other words, participants reproduced the misleading information even when they probably knew the correct answers; suggestibility was not limited to only the most difficult items.

Discussion

Reviewing information makes it more accessible, increasing the probability that the same information will be easily retrievable a short time later (e.g., Berger et al., 1999; Cleary, 2006; Cleary & Reyes, 2009; Cleary & Speckler, 2007). In this case, when participants studied correct knowledge just a few minutes before exposure to misinformation, that correct knowledge remained highly accessible and protected participants from later reproducing the story errors. In contrast, a sizable misinformation effect occurred when participants had studied the relevant facts at a longer delay of 1 week before story reading. Importantly, these same effects occurred even when we restricted the analysis to only those items that participants were most likely to know; the detrimental effects of misinformation exposure were not limited to only the normatively most difficult items.

As was noted above, however, we cannot definitively assess what our individual participants did versus did not know prior to the experiment and, thus, whether the manipulation worked by activating prior knowledge or teaching new knowledge. For items that participants did not know before the experiment, it is unsurprising that they later reproduced the story errors, especially when the story-reading phase took place 1 week after the initial study phase. In this situation, participants likely forgot their new learning from the initial study phase and (1 week later) learned the corresponding misleading answers from the stories, which they then reproduced on the immediate final test. Similar patterns have been demonstrated repeatedly in the eyewitness memory literature (e.g., Loftus, Miller, & Burns, 1978).

In the present data, however, it is likely that *both* prior knowledge and new learning contributed to the effects observed. To separate these possibilities, in Experiments 2a and 2b, we measured preexisting knowledge for each individual

[2] Experiments 2a and 2b showed that, on average, participants had correct stored knowledge for approximately 37 % of the critical items; thus, we estimated that participants in Experiment 1 had stored knowledge for a similar proportion of the items.

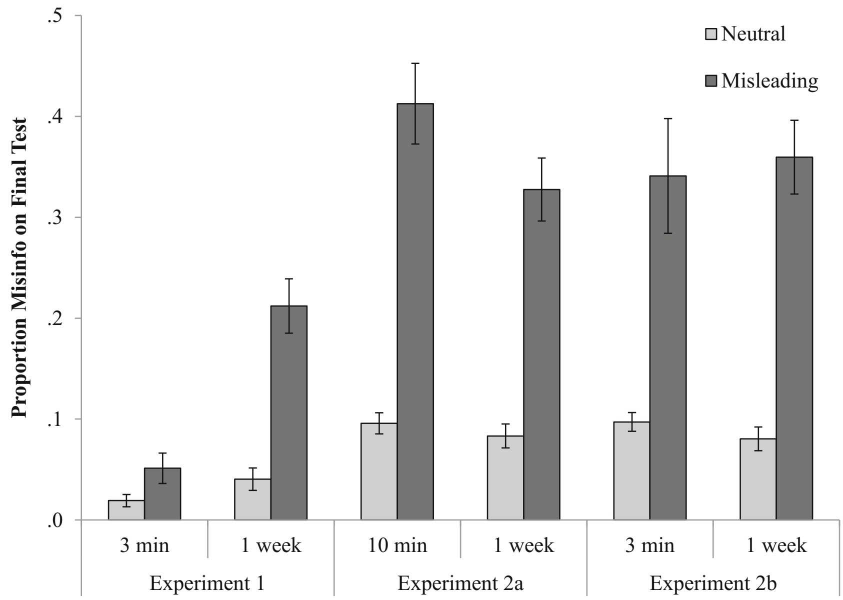

Fig. 1 Misinformation answers produced on the final test in Experiment 1, Experiment 2a, and Experiment 2b after neutral or misleading references in the stories were read. Error bars represent standard errors of the means

participant by having them complete a general knowledge test (instead of a study phase) prior to story-reading.

Experiments 2a and 2b

Because Experiments 2a and 2b were nearly identical, they are reported together. The purpose of both experiments was two-fold: first, to obtain measures of what individual to examine a different way of activating knowledge—namely, retrieval. An abundance of research demonstrates that acts of *retrieving* information from memory produce powerful benefits for memory of the retrieved information (e.g., Carrier & Pashler, 1992; Roediger & Karpicke, 2006a, 2006b). On the basis of this research, we hypothesized that asking participants to retrieve the correct answers from memory would provide a particularly powerful method for heightening the accessibility of that knowledge (see also Ozubko & Fugelsang, 2011; in contrast, see Rapp, 2008). Also consistent with this prediction, Bjork and Bjork (1992) stated in their new theory of disuse that "the act of retrieving an item from memory, and the act of studying an item, both result in increments to that item's retrieval strength and storage strength, but retrieval is the more potent event" (p. 43). In other words, the correct answers should become even more accessible after retrieval than after study, and therefore, retrieval might be expected to provide even greater protection against learning the story errors.

In Experiments 2a and 2b, we modified our procedure to test this possibility. Specifically, rather than asking participants to study the critical questions and their answers at the beginning of the experiment, we instead asked them to *retrieve* the answers to those questions. In both experiments (and similar to Experiment 1), the critical manipulation was the length of the delay between the initial retrieval phase and the story-reading phase. In Experiment 2a, this delay was either 10 min (short-delay condition) or 1 week (long-delay condition). In Experiment 2b, the delays for the short- and long-delay conditions were 3 min and 1 week, respectively. Just as in Experiment 1, we predicted that participants in the short-delay condition would be *more likely* to rely on their prior knowledge and *less likely* to reproduce the misleading story answers; in other words, highly accessible prior knowledge should protect against suggestibility.

Method

Participants

Participants were Duke University undergraduates, run in groups of 1–4, who received course credit or monetary compensation. In Experiment 2a, there were 40 participants in the short-delay condition and 38 participants in the long-delay condition. In Experiment 2b, there were 36 participants in the short-delay condition and 38 participants in the long-delay condition.

Design

As in Experiment 1, the design for both Experiments 2a and 2b was a 2 (delay: short, long) × 2 (fact frame: neutral, misleading) mixed factorial. The length of the delay between the initial retrieval phase and the story-reading phase was the between-subjects variable, and framing of the story references was the within-subjects variable.

Materials

The critical and filler questions and story materials were the same as in Experiment 1. Instead of studying the questions and their correct answers at the beginning of the experiment, participants were now asked to retrieve the answers to those same questions.

Procedure

In both experiments, participants first completed the initial retrieval phase. In Experiment 2a, participants recorded their answers using paper and pencil. The experimenter read each question out loud and paused for up to 20 s between questions, proceeding more quickly if all participants indicated that they were ready to move on. In Experiment 2b, the initial retrieval phase was self-paced, and participants completed it in on a computer in the lab (Jarvis, 2008a, 2008b). The questions appeared one at a time on the screen, and participants typed their answers and pressed the Enter key when they were ready to move to the next question. In both experiments, participants were instructed not to guess and were allowed to answer "I don't know." They rated their confidence in each answer on a 5-point Likert scale.

Participants in the short-delay condition then completed a filler task (word searches and mazes in Experiment 2a and brainteasers in Experiment 2b), which lasted for 10 min in Experiment 2a and 3 min in Experiment 2b. Participants in the long-delay condition were dismissed and reminded to return 1 week later for the second session.

After the short or long delay, participants completed the story-reading phase and the final general knowledge test, which were identical to Experiment 1, before being debriefed.

Results

Data scoring

Responses for both the initial retrieval phase and the final general knowledge test were scored according to the same criteria as in Experiment 1. Two coders scored the responses (Cohen's kappa = .97); a third coder resolved the discrepancies.

Initial retrieval phase

We first examined responses in the initial retrieval phase to determine how much prior knowledge participants had for the critical facts. In Experiment 2a, participants answered an average proportion of .37 of the critical questions correctly, demonstrating prior knowledge for those items; they answered an average of .39 correctly in Experiment 2b. Note that, because the initial retrieval phase was in the form of a short-answer test, these are answers that participants could correctly *produce* (as opposed to recognize) and so could be considered to reflect only strong prior knowledge.

Participants in the short- and long-delay conditions came into the experiments with similar levels of prior knowledge [Experiment 2a, $F(1, 76) = 1.49$, $p = .23$, $\eta^2 = .02$]; Experiment 2b, $F < 1$]. They rarely came into the experiments with misconceptions ($M = .09$ in Experiment 2a and $M = .08$ in Experiment 2b), and this did not differ across the short- and long-delay conditions ($F<1$ for both experiments). In addition, overall, participants were well calibrated in their use of the confidence scale: They were more likely to produce high-confident *correct* responses ($M = 6.25$) than high-confident *incorrect* responses ($M = 0.41$).

Final test performance

As in Experiment 1, our first interest was in how often participants reproduced the story errors on the final general knowledge test (see Fig. 1). Critically, for purposes of comparison to Experiment 1, all of the items were included in this initial analysis (i.e., even those items for which participants had answered the corresponding initial test question incorrectly). In both experiments, there was a significant detrimental effect of misinformation exposure [$F(1, 76) = 117.04$, $MSE = .03$, $p < .01$, $\eta^2 = .60$ in Experiment 2a, and $F(1, 72) = 86.15$, $MSE = .03$, $p < .01$ in Experiment 2b]. That is, even though participants had just attempted to retrieve the answers to the corresponding general knowledge questions, they still reproduced much of the misinformation from the stories ($M = +.28$ after reading the misleading vs. neutral references in Experiment 2a and $M = +.26$ in Experiment 2b). In contrast to our predictions, the fact-frame × delay interaction did not reach significance in either experiment [$F(1, 76) = 1.96$, $MSE = .03$, $p = .17$, $\eta^2 = .01$ in Experiment 2a, and $F < 1$ in Experiment 2b]. In other words, answering the general knowledge questions immediately before the story-reading phase provided no protection (relative to answering those questions 1 week earlier) against the tendency to later reproduce the misleading story answers.

As has already been noted, however, a major goal of Experiments 2a and 2b was to provide a direct measure of what participants knew at the beginning of the experiment, enabling a clearer test of the possibility that heightening the accessibility of *stored knowledge* immediately before misinformation exposure should protect against the later reproduction of that misinformation. Thus, our primary interest was whether participants reproduced the misleading answers from the stories, *even after previously demonstrating knowledge of the correct answers* (see Fig. 2). In both Experiments 2a and 2b, there was a robust suggestibility effect: Participants answered more general knowledge questions with misinformation after reading misleading versus neutral references in the stories [M =+.25 in Experiment 2a and +.22 in Experiment 2b; $F(1, 76) = 53.48$, $MSE = .05$, $p < .001$, $\eta^2 = .40$ for Experiment 2a, and $F(1, 72) = 42.13$, $MSE = .04$, $p < .001$, $\eta^2 = .37$ for Experiment 2b].

Did heightening the accessibility of participants' prior knowledge, by placing the initial retrieval phase just a few minutes before the story-reading phase, reduce this tendency to reproduce the story misinformation? In contrast to Experiment 1, highly accessible knowledge did *not* decrease suggestibility, as compared with less accessible knowledge. In fact, in Experiment 2a, there was a significant interaction between delay and fact-framing, $F(1, 76) = 5.85$, $MSE = .05$, $p = .02$, $\eta^2 = .04$, such that participants in the short-delay condition actually showed a *greater* tendency than did those in the long-delay condition (Ms = .33 vs. .17) to reproduce the misleading answers on the final test. In Experiment 2b, there was no evidence that the size of the misinformation effect differed across the short- ($M = .23$) and long-delay ($M = .22$) groups ($F < 1$ for the delay × fact-frame interaction). Of note is that in both experiments, the same statistical conclusions held when we included only those items for which participants had previously demonstrated high-confidence knowledge [$M_{\text{shortdelay}} = .37$ vs. $M_{\text{longdelay}} = .08$; $F(1, 67) = 8.43$, $MSE = .06$, $p = .01$, $\eta^2 = .09$, for the delay × fact-frame interaction in Experiment 2a, and $M_{\text{shortdelay}} = .07$ vs. $M_{\text{longdelay}} = .11$, $F < 1$ in Experiment 2b].[3] Clearly, completing the initial retrieval phase just a few minutes before the story-reading phase did *not* protect participants from reproducing the story misinformation.

As a second (and not entirely independent) measure of suggestibility, we examined whether exposure to the misleading references decreased participants' abilities to answer correctly on the final test, even though they had demonstrated previously that they knew those same correct answers. Not surprisingly, when participants had previously provided the correct answer and were then exposed to a neutral reference, they were highly likely to produce the same correct answer again on the final test. This finding underscores the relative stability of semantic memory, as highlighted in the introduction (e.g., Bahrick et al., 1975). In both experiments, participants answered fewer final test questions correctly after exposure to the misleading story references [$F(1, 76) = 56.29$, $MSE = .05$, $p < .001$, $\eta^2 = .40$ in Experiment 2a, and $F(1, 72) = 49.60$, $MSE = .05$, $p < .001$, $\eta^2 = .40$ in Experiment 2b; see Table 1]. The critical question was whether this effect differed on the basis of the accessibility of participants' prior knowledge. The results mirrored those for misinformation production: Retrieving one's knowledge immediately before reading misinformation did not prevent the decrease in correct responding following misinformation exposure. In Experiment 2a, there was a significant interaction between delay and fact-framing, $F(1, 76) = 8.15$, $MSE = .05$, $p = .006$, $\eta^2 = .06$, with participants in the short-delay condition showing a *greater* decrease in correct responding after exposure to the misleading references. In Experiment 2b, there was no evidence that the effects of misinformation exposure varied across the delay conditions ($F < 1$). Again, the same statistical conclusions held even when we considered only those answers that participants had originally produced with the highest confidence: a mean decrease in correct responding of .38 vs. .08 for the short versus long delay conditions in Experiment 2a [$F(1, 67) = 13.14$, $MSE = .06$, $p = .001$, $\eta^2 = .12$, for the delay × fact-frame interaction] and a mean decrease in correct responding of .13 vs. .13 in Experiment 2b ($F < 1$ for the corresponding interaction). Participants lost access to some of the correct knowledge they had just demonstrated a few minutes earlier.

[3] On average across experiments, participants produced about 6.25 high-confident correct responses during the initial retrieval phase, a sufficient number to support this analysis.

Discussion

Participants in both experiments showed robust suggestibility, reproducing the misleading answers from the stories even when they had demonstrated prior knowledge of the correct answers. Critically, highly accessible knowledge did not protect against the detrimental effects of misinformation exposure. Instead, these participants were *at least* as likely to reproduce the misleading story answers, in comparison with participants who completed the initial retrieval phase 1 week earlier.

This finding was surprising and in direct contrast to the results of Experiment 1. Some prior work, however, has shown that other manipulations designed to help participants detect errors, such as recalling related information (Rapp, 2008), slowing presentation time (Fazio & Marsh, 2008), and highlighting the errors (Eslick, Fazio, & Marsh, 2011), can fail to decrease (or actually even *increase*) later suggestibility. Here, like highlighting the errors in the stories (Eslick et al., 2011), retrieving the answers may have directed participants' attention toward the relevant story items, actually enhancing the encoding of the misleading answers (Thomas, Bulevich, & Chan, 2010; see also Wissman, Rawson, & Pyc, 2011).

Recent work in the eyewitness memory literature is consistent with this possibility. Chan and colleagues (Chan & Langley, 2011; Chan & LaPaglia, 2011; Chan, Thomas, & Bulevich, 2009; Thomas et al., 2010) have shown that asking participants to take an initial test on details of a recently encoded eyewitness event (e.g., a museum robbery) can paradoxically *increase* the likelihood that they will later reproduce misleading details that contradict the original (and retrieved) information. Although the theoretical mechanisms

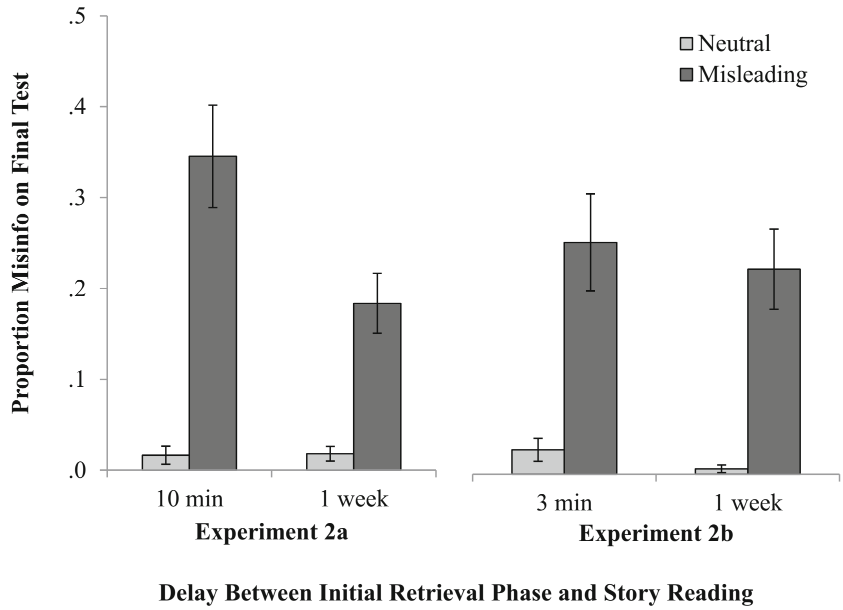

Fig. 2 Misinformation answers produced on the final test in Experiments 2a and 2b after knowledge was demonstrated on the initial test. Error bars represent standard errors of the means

behind this *retrieval-enhanced suggestibility effect* are still up for the debate, this finding fits with the idea that testing enhances future encoding (Chan, Wilford, & Hughes, 2012; Thomas et al., 2010; cf. Arnold & McDermott, 2013; Izawa, 1966, 1970).

A related possibility is that retrieving information from memory causes those memories to enter a temporarily labile state of reconsolidation (e.g., Nader et al., 2000; Walker et al., 2003), allowing new information that is presented during the reconsolidation window to replace or update the original event. In other words, after retrieving the answers to the target general knowledge questions, participants' memories for those answers may have become temporarily malleable, allowing the misinformation answers to become incorporated into the original memory trace of a given fact. Then, on the final general knowledge test at the end of the experiment, participants reported the misleading answer, which likely came to mind easily and was thus interpreted as true (Kelley & Lindsay, 1993).

Table 1 Correct answers on the final general knowledge test, given prior knowledge

	Experiment 2a		Experiment 2b	
	Neutral	Misleading	Neutral	Misleading
Short delay	.96 (.10)	.59 (.38)	.95 (.11)	.67 (.34)
Long delay	.91 (.14)	.75 (.24)	.93 (.10)	.71 (.30)

Note. The short delay was 10 min in Experiment 2a and 3 min in Experiment 2b, and the long delay was 1 week in both experiments. Standard deviations are in parentheses.

General discussion

Exposure to misleading references led participants to reproduce that misinformation, even when they had studied (Experiment 1) or previously retrieved (Experiments 2a and 2b) the correct information. These errors were reproduced after only a single exposure and despite the fact that participants were warned that the stories were fictional and were discouraged from guessing on the final test. In Experiment 1, these effects persisted regardless of item difficulty; participants reproduced the misleading information even for items that we would have expected them to know. Even more surprisingly, after demonstrating correct prior knowledge in Experiments 2a and 2b, participants still reproduced much of the story misinformation.

We explored two methods for protecting participants from incorporating the misinformation into their knowledge bases. One method—requiring participants to *study* the correct answers to the target general knowledge questions before the story-reading phase—successfully protected participants from acquiring many of the story errors, but only when that study phase took place immediately before misinformation exposure. Studying the correct information prior to exposure to misleading content likely had two different benefits: teaching participants new facts, as well as activating information already stored in participants' knowledge bases. The initial study phase likely increased both storage and retrieval strength for all of the correct information (Bjork & Bjork, 1992), providing benefits similar to those observed in the marginal knowledge and tip-of-the-tongue literatures. These benefits were not long-lasting, however, with accessibility of the correct answers decreasing after 1 week.

We expected that retrieving information from memory would provide an even more potent method for heightening the accessibility of correct knowledge (e.g., Ozubko & Fugelsang, 2011) and protecting against the learning of misinformation. In contrast to this prediction, there were sizable misinformation effects in Experiments 2a and 2b, both when participants completed the initial retrieval phase 1 week earlier and even when they completed that phase just a few minutes before the story-reading phase. Although it is surprising that the initial retrieval phase did not prevent later suggestibility, these results are consistent with the idea that acts of retrieval do *more* than simply strengthening the memory trace for the retrieved information. More specifically, retrieval is also thought to affect memory for material occurring *after* the retrieval attempt (e.g., Arnold & McDermott, 2013; Izawa, 1966, 1970). For example, retrieving information from memory may direct people's attention toward relevant items for future encoding (e.g., Chan et al., 2009), it may heighten their curiosity about the correct answer, it may increase their attention to corrective feedback (e.g., Potts & Shanks, 2014), and/or it may render the original memory trace malleable and therefore subject to updating (Nader et al., 2000). To the extent that participants believed that some of the (mis)information in the stories was correct, the initial retrieval phase may have paradoxically enhanced their attention to, and encoding of, those answers. Moreover, because the story information may have been surprising to participants (i.e., it did not match their expectations of the correct answer), the errors may have been especially likely to capture their attention (e.g., Butterfield & Metcalfe, 2001; Fazio & Marsh, 2009). Thus, although the initial retrieval phase should have made participants' correct knowledge more easily retrievable, this benefit was counteracted by the fact that it also seems to have enhanced encoding of the story misinformation.

In considering the larger literature on memory errors, the present data show a similarity between errors of episodic memory and semantic illusions: In both cases, retrieving relevant information does not protect against suggestibility. What is trickier is fitting this finding into the larger pattern of results, wherein sometimes episodic and semantic errors look similar and sometimes they do not (e.g., Marsh & Umanath, in press). For example, preencoding warnings and slower presentation speed reduce errors of episodic memory (Greene, Flynn, & Loftus, 1982; Tousignant, Hall, & Loftus, 1986), but not illusions of knowledge (e.g., Fazio & Marsh, 2008; Marsh & Fazio, 2006). However, repeated exposure and delay have similar effects on the two types of errors. The latter manipulations likely affect participants' memory for the errors, as opposed to affecting participants' abilities to notice the errors while reading. The question here is why a reader would be able to successfully retrieve his or her own knowledge, pay attention to the error, and yet not catch the contradiction between the two (or, if the contradiction is noticed, it is apparently resolved in favor of the story error). These effects happen regardless of the participants' confidence in their initial answers. Accessibility was predicted to increase people's ability to catch contradictions with stored knowledge, but in the case of retrieving stored correct knowledge, it may have, instead, increased encoding without successful monitoring.

Practically, these results suggest that educators should be cautious about using teaching tools that could potentially expose students to misinformation. For example, although films and fictional materials can provide an engaging medium for conveying information about the world, they can also expose learners to erroneous beliefs. Importantly, just because people know relevant information does not mean that they will bring that knowledge to bear. Educators who want to prevent students from incorporating misconceptions into their knowledge base should encourage students to study the correct information before the potential exposure to errors. Having students generate their own knowledge may not be effective, at least in situations like our Experiments 2a and 2b, in which no feedback is given (see Rapp et al., 2013, for a related finding). More generally, these results demonstrate that errors can be easily incorporated into the knowledge base, even when the correct information has recently been studied or retrieved. Just like memories for specific events and experiences, knowledge about the world is quite malleable.

Author Note Portions of this research were presented at the 53rd annual meeting of the Psychonomic Society, Minneapolis, MN. The first and second authors are supported by National Science Foundation Graduate Research Fellowships. This research was supported by a Collaborative Activity award from the James S. McDonnell Foundation's 21st Century Science Initiative in Bridging Brain, Mind and Behavior (EJM). We are grateful to Kathleen Arnold for helpful comments on an earlier version of the manuscript and to Lilly Stiff and Carrie Arndt for their assistance with data collection and coding.

References

Arnold, K. M., & McDermott, K. B. (2013). Test-potentiated learning: Distinguishing between direct and indirect effects of tests. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 39,* 940–945.

Bahrick, H. P. (1984). Semantic memory content in permastore: Fifty years of memory for Spanish learned in school. *Journal of Experimental Psychology: General, 113,* 1–29.

Bahrick, H. P., Bahrick, P. O., & Wittlinger, R. P. (1975). Fifty years of memory for names and faces: A cross-sectional approach. *Journal of Experimental Psychology: General, 104,* 54–75.

Bahrick, H. P., & Hall, L. K. (1991). Lifetime maintenance of high school mathematics content. *Journal of Experiment Psychology: General, 120,* 20–33.

Bahrick, H. P., & Phelps, E. (1988). The maintenance of marginal knowledge.

Berger, S. A., Hall, L. K., & Bahrick, H. P. (1999). Stabilizing access to marginal and submarginal knowledge. *Journal of Experimental Psychology: Applied, 5,* 438–447.

Bjork, R. A., & Bjork, E. L. (1992). A new theory of disuse and an old theory of stimulus fluctuation. In A. Healy, S. Kosslyn, & R. Shiffrin (Eds.), *From learning processes to cognitive processes: Essays in honor of William K. Estes* (Vol. 2, pp. 35–67). Hillsdale, NJ: Erlbaum.

Bottoms, H. C., Eslick, A. N., & Marsh, E. J. (2010). Memory and the Moses Illusion: Failures to detect contradictions with stored knowledge yield negative memorial consequences. *Memory, 18,* 670–678.

Brown, A. S. (2012). *The tip of the tongue state: Essays in cognitive psychology* (Vol. 1), Psychology Press.

Butterfield, B., & Metcalfe, J. (2001). Errors committed with high confidence are hypercorrected. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27,* 1491–1494.

Carrier, M., & Pashler, H. (1992). The influence of retrieval on retention. *Memory and Cognition, 20,* 632–642.

Chan, J. C. K., & Langley, M. (2011). Paradoxical effects of testing: Retrieval enhances both accurate recall and suggestibility in eyewitnesses. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 37,* 248–255.

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

Chan, J. C. K., Wilford, M. M., & Hughes, K. L. (2012). Retrieval can increase or decrease suggestibility depending on how memory is tested: The importance of source complexity. *Journal of Memory and Language, 67,* 78–85.

Cleary, A. M. (2006). Relating familiarity-based recognition and the tip-of-the-tongue phenomenon: Detecting a word's recency in the absence of access to the word. *Memory & Cognition, 34,* 804–816.

Cleary, A. M., & Reyes, N. L. (2009). Scene recognition without identification. *Acta Psychologica, 131,* 53–62.

Cleary, A. M., & Specker, L. E. (2007). Recognition without face identification. *Memory & Cognition, 35,* 1610–1619.

Eslick, A. N., Fazio, L. K., & Marsh, E. J. (2011). Ironic effects of drawing attention to story errors. *Memory, 19,* 184–191.

Estes, W. K. (1955a). Statistical theory of spontaneous recovery and regression. *Psychological Review, 62,* 145–154.

Estes, W. K. (1955b). Statistical theory of distributional phenomena in learning. *Psychological Review, 62,* 369–377.

Fazio, L. K., Barber, S. J., Rajaram, S., Ornstein, P. A., & Marsh, E. J. (2013). Creating illusions of knowledge: Learning errors that contradict prior knowledge. *Journal of Experimental Psychology: General.*

Fazio, L. K., & Marsh, E. J. (2008). Slowing presentation time increases, rather than decreases, errors learned from fictional stories. *Psychonomic Bulletin & Review, 15,* 180–185.

Fazio, L. K., & Marsh, E. J. (2009). Surprising feedback improves later memory. *Psychonomic Bulletin & Review, 16*, 88-92.

Greene, E., Flynn, M. S., & Loftus, E. F. (1982). Inducing resistance to misleading information. *Journal of Verbal Learning and Verbal Behavior, 21,* 207–219.

Izawa, C. (1966). Reinforcement-test sequences in paired-associate learning. *Psychological Reports, 18,* 879–919. doi:10.2466/pr0.1966.18.3.879

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

Jarvis, B. G. (2008a). *DirectRT (Version 2008.1.0.13*) [Computer software]. New York, NY: Empirisoft Corporation.

Jarvis, B. G. (2008b). *MediaLab (Version 2008.1.33*) [Computer software]. New York, NY: Empirisoft Corporation.

Kelley, C. M., & Lindsay, D. S. (1993). Remembering mistaken for knowing: Ease of retrieval as a basis for confidence in answers to general knowledge questions. *Journal of Memory and Language, 32,* 1–24.

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

Loftus, E. F., & Pickrell, J. E. (1995). The formation of false memories. *Psychiatric Annals, 25,* 720–725.

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, 41*, 19–31.

Marsh, E. J. (2004). Story stimuli for creating false beliefs about the world. *Behavior Research Methods, Instruments, & Computers, 36*, 650–655.

Marsh, E. J., & Fazio, L. K. (2006). Learning errors from fiction: Difficulties in reducing reliance on fictional stories. *Memory & Cognition, 34*, 1140–1149

Marsh, E. J., Meade, M. L., & Roediger, H. L., III. (2003). Learning facts from fiction. *Journal of Memory and Language, 49,* 519–536.

Marsh, E. J. & Umanath, S. (in press). Knowledge neglect: Failures to notice contradictions with stored knowledge. In D. N. Rapp, & J. Braasch (Eds.) *Processing Inaccurate Information: Theoretical and Applied Perspectives from Cognitive Science and the Educational Sciences.*

Myers, J. L., O'Brien, E. J., Balota, D. A., & Toyofuku, M. L. (1984). Memory search without interference: The role of integration. *Cognitive Psychology, 16,* 217–242.

Nelson, T. O., & Narens, L. (1980). Norms of 300 general-information questions: Accuracy of recall, latency of recall, and felling-of-knowing ratings. *Journal of Verbal Learning and Verbal Behavior, 19*, 338–368.

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

Ozubko, J. D., & Fugelsang, J. (2011). Remembering makes evidence compelling: Retrieval from memory can give rise to the illusion of truth. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 37,* 270–276.

Prentice, D. A., Gerrig, R. J., & Bailis, D. S. (1997). What readers bring to the processing of fictional texts. *Psychonomic Bulletin & Review, 4,* 416–420.

Potts, R., & Shanks, D. R. (2014). The benefit of generating errors during learning. *Journal of Experimental Psychology: General, 143*, 644–667.

Rapp, D. N., Hinze, S. R., Kohlhepp, K., & Ryskin, R. A. (2013). Reducing reliance on inaccurate information. *Memory and Cognition*.

Rapp, D. N. (2008). How do readers handle incorrect information during reading? *Memory & Cognition, 36*, 688–701.

Rastle, K. G., & Burke, D. M. (1996). Priming the tip of the tongue: Effects of prior processing on word retrieval in young and older adults. *Journal of Memory & Language, 35,* 586–605.

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

Roediger, H. L., & Karpicke, J. D. (2006b). Test-enhanced learning: Taking memory tests improves long-term retention. *Psychological Science, 17,* 249–255.

Tauber, S. K., Dunlosky, J., Rawson, K. A., Rhodes, M. G., & Siltzman, D. M. (2013). General knowledge norms: Updated and expanded from the Nelson and Narens (1980) norms. *Behavioral Research Methods, 45*, 1115–1143.

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.

Tousignant, J. P., Hall, D., & Loftus, E. F. (1986). Discrepancy detection and vulnerability to misleading post-event information. *Memory & Cognition, 14,* 329–338.

Tulving, E. (1983). Echphoric processes in episodic memory. Phil. Trans. R. Soc. Lond. B., *Biological Science, 302*(1110), 361–370.

Tulving, E., & Pearlstone, Z. (1966). Availability versus accessibility of information in memory for words. *Journal of Verbal Learning and Verbal Behavior, 5,* 381–391.

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

Wheeler, S. C., Green, M. C., & Brock, T. C. (1999). Fictional narratives change beliefs: Replications of Prentice, Gerrig, & Bailis (1997) with mixed corroboration. *Psychonomic Bulletin and Review, 6,* 136–141.

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