



## Learning facts from fiction<sup>☆</sup>

Elizabeth J. Marsh,<sup>\*</sup> Michelle L. Meade, and Henry L. Roediger III

*Washington University in St. Louis, MO 63130, USA*

Received 26 February 2003; revision received 25 June 2003

### Abstract

People's knowledge about the world comes from many sources, including fictional ones such as movies and novels. In three experiments, we investigated how people learn and integrate information from fictional sources with their general world knowledge. Subjects read a series of short stories that contained information about the real world. After a short delay, all participants took a general knowledge test. Subjects did indeed use information from the stories to answer general knowledge questions. Prior reading of facts boosted participants' abilities to produce both obscure and better-known facts, and the effect held for both correct and incorrect facts (misinformation). Repeated reading of the stories increased the effect. After a delay of one week, effects of story exposure were strongest for items that also had been tested in the first session. Subjects were aware of using story information, but interestingly, story exposure also increased belief that the facts had been known prior to the experiment, even for misinformation answers that were rarely produced without story reading.

© 2003 Elsevier Inc. All rights reserved.

*Keywords:* Illusory truth; Fiction; False memory; Source monitoring

### Introduction

People learn information about the world from a multitude of sources: other people, newspapers, textbooks, classes, museums, and so on. While encyclopedias, non-fiction books, documentaries, and other such sources are designed to teach, learning may also

result from exposure to non-educational sources that happen to contain information about the world. Fictional sources such as television sitcoms, movies, novels, short stories, and even comic strips often occur in familiar political, geographical, and historical contexts. As such, fiction is potentially a source of information about the world. Clearly, educators believe that students can benefit from exposure to fictional materials (e.g., Dubeck, Bruce, Schmuckler, Moshier, & Boss, 1990; Smith, 1993; Storey, 1982). For example, the curriculum in a history course might include Dostoyevsky's novels as a vehicle for learning about Russian culture and history. Not all fictional sources, however, are as accurate as non-fictional ones. For example, not all science fiction novels and movies are good sources of information about physics and other sciences. The current research program is generally concerned with these issues, including the extent to which students rely on fictional sources when answering general knowledge questions, and their awareness of this reliance.

<sup>☆</sup> We thank Patrick Dolan, Keith Payne, and Dave Balota for discussions of theory and data. Colleen Kelley and Alan Brown provided many useful comments on a draft of the manuscript. We are grateful to Kristina Olson, Rebecca Roediger, and Animesh Tandon for help with data collection and input. This research was supported by an NRSA postdoctoral fellowship from the National Institute of Mental Health (#1F32MH12567) to the first author, and Grant RO1 AG17481-01A1 from the National Institutes of Aging to the third author.

<sup>\*</sup> Corresponding author. Current address: Psychological & Brain Sciences, Duke University, 9 Flowers Drive, Box 90086, Durham, NC 27708-0086, USA.

*E-mail address:* [emarsh@psych.duke.edu](mailto:emarsh@psych.duke.edu) (E.J. Marsh).

When thinking about these issues, it is important to consider the theoretical distinction in the text processing literature between *integration* and *compartmentalization* (e.g., Potts & Peterson, 1985). Integration of facts from fiction would mean that readers link these facts to pre-existing world knowledge. In its strongest form, integrated 'fictional' facts would be represented in the same fashion as other related world knowledge, without retaining links to their fictional sources. In contrast, compartmentalization would occur if readers represented these 'fictional' facts in memory as separate from the rest of their general world knowledge. Such compartmentalization might occur for a number of reasons, such as the belief that the fictional source was not credible or a failure to make a connection between the facts in the fictional world and the real world. More recently, a *hybrid* position has been suggested, in which facts learned from fiction are associated with related world knowledge, but also retain associations with the fictional source (Gerrig & Prentice, 1991; Prentice & Gerrig, 1999).

#### *Evidence for integration*

The question is whether students integrate information from fictional sources with their general world knowledge, just as they do after learning facts in a class or from a textbook. Eventually, the knowledge retained from a class is integrated to the extent that students shift from labeling these facts as 'remembered' to 'known' (Conway, Gardiner, Perfect, Anderson, & Cohen, 1997). In conceptual implicit memory studies, even a single reading of correct and incorrect answers (e.g., 'Hickock') in a list of words increases their later production as answers to general knowledge questions (e.g., 'What was Buffalo Bill's last name?'; Blaxton, 1989; Kelley & Lindsay, 1993; Thapar & Rouder, 2001). If reading fiction is like reading a list of correct and incorrect answers, or listening to a lecture or reading a textbook, we would predict that students would also draw on fiction as a source of information about the real world.

Educators' use of fiction in the classroom suggests that they believe students integrate information from fictional sources with the rest of their world knowledge. Several experimental results support this view. For example, in Lewis and Anderson's (1976) classic study, subjects learned varying numbers of fantasy facts about real people, and then later verified true facts. The more fantasy facts subjects had studied, (e.g., Napoleon Bonaparte was from India), the slower they were to respond to true related facts (e.g., Napoleon Bonaparte was an emperor). Fantasy facts interfered with access to prior knowledge, suggesting the two were integrated in memory (see also Peterson & Potts, 1982). Gerrig and Prentice (1991) extended these results, showing that fantasy facts embedded in much longer stories can also

slow later rejection of false facts. In some circumstances, reading fantasy facts embedded in stories can even change subjects' ratings of agreement with story facts (Prentice, Gerrig, & Bailis, 1997; Strange & Leung, 1999; Wheeler, Green, & Brock, 1999).

#### *Limits on integration of facts from fiction*

Students may be reluctant to treat fiction as fact; in one survey, Yale undergraduates agreed that authors of fiction sometimes invent facts for storyline purposes (Prentice & Gerrig, 1999). In numerous other domains, subjects avoid relying on low-credibility sources (e.g., Hoffman, Granhag, See, & Loftus, 2001; Hovland, Lumsdaine, & Sheffield, 1949). For example, in eyewitness studies, suggestibility is reduced when subjects know the post-event communication was produced by a less credible source such as a naïve interviewer (Smith & Ellsworth, 1987) or a defense lawyer (Dodd & Bradshaw, 1980). And once eyewitnesses notice errors in a narrative, there is often a spillover effect in which blatant misinformation serves as a warning and reduces suggestibility (Loftus, 1979). Readers of fiction do appear to be monitoring the text for accuracy; for example, there is less interference from wrong 'facts' that contradict the current state of the world (Gerrig & Prentice, 1991), and subjects who are better able to evaluate story information are less influenced by the stories (e.g., Prentice et al., 1997; Wheeler et al., 1999). Thus, there may be limits on which information from fiction is integrated with related world knowledge.

Other data suggest that facts learned from fiction retain links to their fictional sources. If 'fictional' facts were fully integrated, then it would not matter whether or not the test reinstated the story context—but type of test matters. For example, Lewis and Anderson's (1976) subjects were faster to verify true facts when the test did not contain any of the studied fantasy facts (a pure test) than when the test contained a combination of true and fantasy facts (a mixed test). This finding suggests that some form of source information was retained, and accordingly subjects' behavior was affected by the two different sources in the mixed test. Similarly, when text had been labeled as 'fiction,' subjects were faster to verify story facts when they were embedded in a block of story items than when they were tested in a block of non-story items (Gerrig & Prentice, 1991; Potts & Peterson, 1985; Potts, St. John, & Kirson, 1989). In contrast, labeling the same information as 'fact' led to items being verified more quickly when in a block of non-story items (Potts & Peterson, 1985; Potts et al., 1989).

Linking to the story source may be more likely for facts that seem suspicious, that contradict general world knowledge. When Peterson and Potts (1982) switched the fantasy facts to obscure ones (rather than facts that contradicted most people's general knowledge), the

pure-mixed test difference was lessened. When the fantasy facts were less blatantly false, subjects likely felt less of a need to compartmentalize the information and simply integrated the new facts with prior knowledge. Similarly, Gerrig and Prentice (1991) found that having read blatantly contradictory items (e.g., that Geraldine Ferraro was the Vice President of the USA) did not slow verification of the true state of the world.

### *The current research program*

In the current research program, we sought to further understand the extent to which facts from fiction are integrated versus compartmentalized. In three experiments, subjects read a series of short stories that contained peripheral references to facts about the real world. Some of these facts corresponded to questions on a later general knowledge test. The basic design of all experiments was very similar to the eyewitness post-event information paradigm (e.g., Loftus, Miller, & Burns, 1978), except that our subjects' original learning occurred (if at all) outside of the laboratory.

The first major goal of the research program was to examine whether subjects would produce facts from fiction on a later test of world knowledge. If facts from fiction are strongly associated with their fictional source, then subjects may not use them on the final test. In particular, in Experiments 2 and 3, we investigated the effects of having read misinformation in the stories. Facts from fiction may be associated with one's world knowledge enough to slow retrieval or slightly shift beliefs, but that does not mean they will be produced as hard facts on a later test. For example, in Gerrig and Prentice's (1991) experiment, reading false facts affected reaction times but did not lead to high error rates. Even when the false fact had been established as a competitor in memory, it was not nearly the strength of the correct answer.

A second goal of the research was to assess whether subjects knew when they were using story facts to answer general knowledge questions. Previously, links to the story source have been inferred, from data points such as Peterson and Potts' (1982) finding that subjects were slower to verify fantasy facts on a test that contained facts from both fantasy and true facts. In contrast, when Green and Brock (2000) asked subjects to remember whether a narrative had been labeled as fact or fiction, many subjects made errors (e.g., 1/3 of subjects made errors in Experiment 1). Thus, one object of the current research was to directly evaluate people's awareness of their reliance on fictional sources. Subjects' source judgments are informative about the underlying representations. That is, if facts from fiction were completely integrated with related world knowledge, we would not expect subjects to be able identify which answers had come from the stories. We would only expect

them to be able to identify the story source if the facts were represented with associations to that source (e.g., the hybrid or compartmentalization positions).

A second source judgment was also of interest: that of pre-experimental knowledge. People often fall prey to the knew-it-all-along bias (Fischhoff, 1977; Wood, 1978), in which subjects are unable to estimate what they would have known had they not been told the correct answer to a question. Similarly, subjects misattribute ease of retrieval at test (due to prior study) to pre-experimental knowledge (Begg, Robertson, Gruppisop, Anas, & Needham, 1996) and to high confidence in responses (Kelley & Lindsay, 1993). These results suggest that if story reading increases the ease with which answers come to mind at test, subjects may think they knew these answers from sources other than story reading. This position is consistent with a representation of 'fictional' facts that is integrated with other pre-experimental knowledge. On the other hand, an illusion of prior knowledge may not occur if subjects remember the story source, since subjects know that they should not believe everything they read in a fictional source (Gerrig & Prentice, 1991). Put another way, if facts from fiction are compartmentalized, kept separate from related world knowledge, we would not expect them to increase estimates of pre-experimental knowledge.

Thus, the combination of the two source judgments will help elucidate the underlying representation of facts learned from fiction. Integration would be supported by the forgetting of the story source combined with general world knowledge attributions. Compartmentalization would be supported by memory for the story source without a corresponding increase in world knowledge attributions. The hybrid position would be supported by both memory for story source and an increase in world knowledge attributions.

In all experiments, we included two manipulations aimed at understanding *why* students might rely on fictional sources. First, we manipulated the familiarity of the facts. We thought subjects might better remember the story source for unfamiliar facts, since these would be less likely to be associated with other pre-experimental sources. In Experiments 2 and 3, we included misinformation, and we reasoned that subjects would be more likely to notice misinformation for commonly known facts. Subjects should be less suggestible on well-known items if they do in fact selectively use story information (e.g., Gerrig & Prentice, 1991).

Second, we manipulated the number of times the story was read. Reading a story twice should boost subjects' ability to remember the source of the read facts (e.g., Johnson, Raye, & Durso, 1980). Reading a story twice should also increase the likelihood of the facts coming to mind at test (Jacoby, Jones, & Dolan, 1998), potentially increasing story reliance. In addition, since subjects are presumably less involved in the plot during

the second reading of the same story, they should be able to allocate more attention to evaluating background information. Thus, if subjects are actively monitoring the stories, they should show less suggestibility after reading it twice. If subjects do not actively monitor the stories, then reading the same story twice should increase suggestibility as the repeatedly read facts will come easily to mind during the general knowledge test.

To summarize: All three experiments involved story reading, followed by a delay, and then a final test of general world knowledge. Of interest was whether our academically oriented subjects would rely on fictional sources when answering easy and hard general knowledge questions. In Experiment 1, the stories were a potential source of correct information about the world. In Experiments 2 and 3, the stories were a potential source of both correct and incorrect information about the world. In all experiments, we manipulated prior knowledge of the facts and number of story readings. We also assessed subjects' awareness of the story sources, and subjects' beliefs of prior knowledge of story facts. In Experiment 1, we manipulated when subjects made these source judgments, either immediately after answering each general knowledge question or retrospectively after they had answered all questions. Of interest was whether requiring immediate source judgments would increase the salience of the stories throughout the test, potentially causing subjects to perform differently than when they were less aware of the connection between the two phases of the experiment.

## Experiment 1

### Method

#### Subjects

Twenty eight Washington University undergraduates participated in the experiment for partial fulfillment of a course requirement. Four subjects were excluded from the analyses because they reported finishing fewer than 7 of the 9 stories. Thus, 24 subjects were included in the analyses that follow.

#### Design

The experiment consisted of a 2 (question difficulty: easy or hard)  $\times$  2 (fact framing: correct or neutral)  $\times$  3 (number of story readings: 0, 1, or 2)  $\times$  2 (timing of source judgments: immediate or retrospective) mixed design. All variables were manipulated within-subject except for the manipulation of the timing of the source judgments. All variables were counterbalanced across subjects. Of interest was the effect of these manipulations on two dependent variables: proportion of critical test questions answered correctly, and the corresponding source attributions for those answers.

### Materials

Nine fictional short stories were created. Each story was 2–3 single spaced pages and contained characters, dialogue, and plot. Stories covered diverse topics such as an outdoor expedition in Alaska, an art thief, a medical student's first day of class, a scientific inventor, letters written during the Civil War, an older couple's 50th wedding anniversary, a cruise, a girl scout troop, and a planetarium. Each story referred to eight facts from the Nelson and Narens (1980) general knowledge norms. Four of these were defined as high prior knowledge and corresponded to easy test questions; on average, 70% of students in Nelson and Narens' study answered these questions correctly. Four facts in each story were defined as low prior knowledge and corresponded to difficult test questions; on average, 15% of students in Nelson and Narens' study answered these questions correctly.

Within each story, half the facts were framed neutrally whereas the others were framed correctly. A Neutral framing involved a reference to the later general knowledge question without providing the answer to that question. A Correct framing included both the correct answer and a reference to the later general knowledge question. For example, sextant is the correct answer to the general knowledge question "What is the name of the navigation instrument used at sea to plot position by the stars?" The dialogue in the Correct version read "This here, this is a sextant and it's the main tool used at sea to navigate via the stars" whereas the Neutral version read "This here, this is the main tool used at sea to navigate via the stars."

Twelve different story booklets were constructed to counterbalance fact framing (neutral, correct), number of story readings (0, 1, 2), and story order. Two different story orders were used; these orders were random with the constraint that same story was never read twice in a row. In addition, six different question booklets were constructed to match the story booklets (fact framing did not affect the question booklets). Each question booklet contained nine pages (corresponding to the 9 stories) with 4 questions. The first question was always "Did you finish reading the story?" The remaining 3 questions were about plot details; different questions were used for the second reading of the stories.

The general knowledge test contained 164 questions, 72 of which represent the dependent measures in the study. These 72 critical questions corresponded to the 8 critical items from each of the 9 stories; as described already, half were easy questions and half were hard questions, and they had been read in the stories in either a Neutral or Correct framing. The remaining items (fillers) were selected from the Nelson and Narens (1980) norms so as to represent a range of difficulty and to avoid overlap with the critical questions. The first twenty questions on the test were easy filler items.

The questions corresponding to story facts were then randomly intermixed with filler items and tested so that at least 4 questions separated references to the same story.

Each general knowledge question was in a cued recall format, with an answer space for recording the answer. Two small boxes followed each answer space, one labeled GK (for general knowledge) and the other S (for stories). These abbreviations were used so that subjects in the retrospective source condition would not realize the connection of the general world knowledge test to the stories until the appropriate time.

#### Procedure

Subjects were tested in small groups ranging from 1 to 5 people. The experimenter introduced the study as a reasoning experiment. There were three phases: story reading, a filler task, and the test of general world knowledge.

In Phase I, subjects worked through a story booklet and a question booklet. Participants were told to read the short stories carefully as after each story they would answer three comprehension questions. As established in pre-testing, subjects were given up to 5 min to read each story; the experimenter verbally warned subjects at the 3.5 min mark. Participants were instructed to read each story only once, and subjects who finished reading early were instructed to wait quietly. Subjects were told that some stories would be repeated, but that if so the comprehension questions would be different. Subjects completed a total of 9 story-question cycles.

In Phase II, the filler phase, subjects solved a series of visuo-spatial brain-teasers for 7 min. In keeping with the experiment's cover story, these were labeled as non-verbal reasoning tasks.

In Phase III, the test phase, subjects took the 164-item test of general world knowledge. Subjects were instructed to answer the questions in sequential order. They were told to avoid guessing, to only answer the questions they knew, and to draw a line in the answer space if they could not answer the question. In the Immediate condition, the subjects received their source instructions at the beginning of the test and made each source judgment immediately after answering each question. In the Retrospective condition, subjects received their source instructions after they had completed the test of general world knowledge, and they then worked through the entire test a second time to make their source judgments. Regardless of the timing of the source judgments, all participants made two judgments for each answer. First, they decided whether or not they knew their answer based on their general knowledge, meaning that they could have answered the question before the experiment. Second, they decided whether or not their answer had been presented in the stories. For both of these decisions, subjects made yes–no decisions and recorded their

answer in the appropriate boxes. It was made clear to the subjects that they could answer 'yes' to both questions, 'no' to both, or use either combination of 'yes' and 'no.'

Subjects in the Immediate condition were given 32 min to complete the test of general world knowledge. Subjects in the Retrospective condition were given 20 min to record their answers to the questions, and then 12 min to complete the source judgments. Subjects in the Retrospective condition were told not to change any of their answers; to ensure this, they made their source judgments in a different colored ink. To help subjects in both conditions pace their performance, the experimenter periodically let subjects know how time was passing.

Subjects were thanked for their participation and fully debriefed. They were asked not to discuss the experiment with other potential subjects.

#### Results

All results were significant at the .05 level unless otherwise noted.

#### Correct answers

A 2 (question difficulty: easy or hard)  $\times$  2 (fact framing: neutral or correct)  $\times$  3 (number of story readings: 0, 1, or 2)  $\times$  2 (timing of source judgments: immediate or retrospective) ANOVA was computed on proportion of questions answered correctly. There was no main effect of timing of source test; subjects did not answer significantly more questions correctly when all source judgments were made retrospectively after the entire test ( $M = 56\%$ ) than when source judgments were made immediately after each question ( $M = 48\%$ ). The timing of source judgments did not interact with any of the other factors (all  $F$ 's  $< 1$ ). Thus, for simplicity, the data shown in the top panel of Table 1 are collapsed over the timing of source judgments. As expected, subjects correctly answered more easy questions than hard,  $F(1, 22) = 567.87$ ,  $MSE = .03$ . More questions were correctly answered after having read the correct facts in the stories,  $F(1, 22) = 64.20$ ,  $MSE = .03$ , and also with additional story readings,  $F(2, 44) = 17.91$ ,  $MSE = .03$ . There was no interaction between question difficulty and number of story readings.

The key analysis was whether reading the stories increased performance after reading the correct answers, but not after reading the neutral question frames. Supporting our hypotheses, the interaction between fact framing and number of story readings was significant,  $F(2, 44) = 11.57$ ,  $MSE = .03$ , and this effect was not dependent on question difficulty,  $F(2, 44) = 1.48$ ,  $MSE = .03$ . If subjects had not read the relevant story, there was no difference between neutral and correct items for easy,  $t(23) = 1.09$ ,  $SEM = .05$ , or hard questions,  $t < 1$ . One story reading led to a significant difference

Table 1

Mean proportion of questions answered correctly, as a function of question difficulty, number of story readings, and experimental condition

	Easy questions			Hard questions		
	Number of story readings			Number of story readings		
	Zero	One	Two	Zero	One	Two
Exp. 1						
Correct	.66	.86	.85	.22	.45	.51
Neutral	.60	.65	.68	.18	.17	.18
Exp. 2						
Correct	.61	.77	.84	.10	.37	.57
Neutral	.69	.60	.57	.14	.19	.13
Misleading	.62	.53	.37	.19	.11	.11
Exp. 3						
Immediate						
Correct	.63	.82	.80	.23	.36	.49
Misleading	.67	.56	.44	.21	.13	.19
Delay, not tested						
Correct	.66	.78	.75	.23	.27	.27
Misleading	.69	.67	.56	.15	.20	.21
Delay, tested						
Correct	.69	.81	.81	.19	.35	.36
Misleading	.72	.61	.56	.19	.18	.17

between neutral and correct items for both easy,  $t(23) = 4.73$ ,  $SEM = .04$ , and hard questions,  $t(23) = 5.0$ ,  $SEM = .06$ . A similar difference between neutral and correct items was observed after two story readings.

#### Source attributions

How aware were subjects that the facts had appeared in the stories? Table 2 shows the conditional probabilities of making a story attribution, given that the correct answer was produced. Note that not all subjects are included in each mean, as many subjects did not correctly answer questions in all cells (especially hard questions). However, the same pattern of results was obtained when the dependent measure was the joint probability of answering correctly and attributing an item to the story (and this analysis included all subjects).

Subjects rarely thought they had read answers that were embedded in unread stories. They were also fairly good at knowing that the answers to neutral frames had not been in the stories, although this false alarm rate was higher than for unread stories. Subjects were quite aware of when they had read the answers to both easy and hard questions in the story. Separate 2 (fact framing)  $\times$  3 (number of story readings) ANOVAs were computed on the conditional probabilities for easy and hard questions. The interaction was significant for both easy,  $F(2, 44) = 30.42$ ,  $MSE = .04$ , and hard questions,  $F(2, 10) = 18.71$ ,  $MSE = .05$ . That is, story reading increased story attributions more for correct than for neutral frames.

An additional interesting pattern was observed in the general world knowledge attributions. Fig. 1 shows the joint probabilities of having correctly answered a question *and* attributed that answer to prior knowledge (the sum of correct answers attributed to 'both story and general knowledge' and 'general knowledge only'). Having previously read the answers in the stories led to an increase in correct answers that were attributed to general world knowledge. This difference was significant for both easy,  $t(23) = 4.56$ ,  $SEM = .03$ , and hard questions,  $t(23) = 4.16$ ,  $SEM = .04$ . Subjects produced more correct answers attributed to prior knowledge after reading easy ( $M = .76$ ) and hard answers ( $M = .34$ ), as compared to when they had not read the stories or had read only the neutral frames (means of .62 and .19 for easy and hard questions, respectively).

#### Discussion of Experiment 1

Subjects' behavior was quite adaptive; their performance on the test of general world knowledge improved because of their reliance on fictional sources. Most surprising were the source data. First, subjects were quite aware that they were using story information to answer the questions, even though the stories contained only correct information (and thus there was no obvious need to remember the story source). Facts retained links to the story source. However, they also appeared to be integrated with pre-experimental

Table 2

Proportion of correct answers that subjects remembered reading in the stories, as a function of question ease, number of story readings, and experiment condition

	Easy questions			Hard questions		
	Number of story readings			Number of story readings		
	Zero	One	Two	Zero	One	Two
Exp. 1						
Correct	.08	.80	.83	.00	.83	.90
Neutral	.09	.29	.24	.11	.02	.11
Exp. 2						
Correct	.03	.60	.83	.00	.79	.92
Neutral	.06	.32	.21	.04	.30	.18
Misleading	.02	.49	.35	.00	.35	.44
Exp. 3						
Delay, not tested						
Correct	.01	.20	.31	.04	.31	.46
Misleading	.02	.13	.08	.05	.24	.26
Delay, tested						
Correct	.04	.49	.66	.03	.69	.75
Misleading	.08	.29	.27	.05	.28	.31

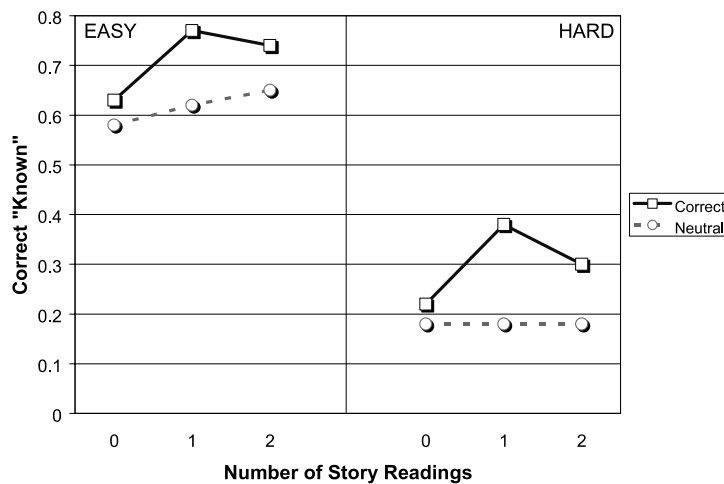


Fig. 1. Joint probability of answering a question correctly and saying 'yes' the answer was known prior to the experiment (Exp. 1).

knowledge, as subjects often claimed that they had known this information *before* the experiment. Begg et al. (1996) observed a similar illusion of knowing after subjects read a list of facts. In neither our experiment nor Begg's, however, can we separate whether subjects were over-estimating their prior knowledge or whether the stories served to remind them of what they already knew. In Experiment 2, we examined these possibilities via source attributions for misinformation answers; misinformation answers should not have been 'known' prior to the experiment. We were of course also interested in the rate of misinformation production as a function of question difficulty and number of story readings.

## Experiment 2

### Method

#### Subjects

Thirty eight undergraduates participated in the experiment for partial fulfillment of a course requirement. Two subjects were eliminated because they did not finish reading at least 7 of the 9 stories, leaving 36 subjects in the analyses.

#### Design

The experiment consisted of a 2 (question difficulty: easy or hard)  $\times$  3 (framing of facts: correct, neutral, or

misleading)  $\times$  3 (number of story readings: 0, 1, or 2) within-subjects design. The dependent variables were subjects' responses on critical test questions, and the corresponding source attributions for those items.

### Materials

The 9 stories from Experiment 1 were modified so that each referred to six critical facts (three corresponding to easy questions and three corresponding to hard questions). The six facts were evenly split among Neutral, Correct, and Misleading framings. Correct and Neutral framings were the same as in Experiment 1; Misleading versions suggested incorrect answers. For example, if the Correct version read "This here, this is a sextant and it's the main tool used at sea to navigate via the stars," the Misleading version read "This here, this is a compass and it's the main tool used at sea to navigate via the stars." Fact framing was counterbalanced across subjects.

Additional materials were created for Experiment 2 to probe subjects' awareness of the misinformation and to aid in debriefing. A questionnaire contained increasingly specific open-ended questions aimed at assessing participants' awareness of the experiment's purpose, beginning with the question "What did you think the experiment was about?" and ending by requiring subjects to list specific pieces of misinformation they had noticed. In addition, a debriefing sheet listed the correct answers to the facts that had been presented in misleading frames. Using a 3-point scale, subjects were asked to rate how surprising they found each of the corrected statements.

### Procedure

Procedurally, the first three phases were the same as in Experiment 1. Experiment 2 contained a fourth phase in which subjects completed the questionnaire assessing their awareness of the experiment's purpose and the presence of misinformation. As part of the debriefing, subjects rated how surprising they found the corrected versions of facts on which they had been misled.

### Results

#### Correct answers

A 2 (question difficulty: easy or hard)  $\times$  3 (fact framing: correct, neutral, or misleading)  $\times$  3 (number of story readings: 0, 1, or 2) ANOVA was conducted on proportion of correct answers. The data are shown in the middle portion of Table 1. Replicating Experiment 1, subjects correctly answered more easy questions and more questions corresponding to correctly framed facts. All of two-way interactions were significant, but the three-way was not ( $F < 1$ ). Fact framing had a bigger effect on subjects' ability to answer easy questions than hard,  $F(2, 70) = 3.26$ ,  $MSE = .05$ . The number of story readings had a bigger effect on subjects' ability to answer

hard questions than easy,  $F(2, 70) = 6.85$ ,  $MSE = .06$ , although this effect was qualified by fact framing.

Most critically, the effect of fact framing was dependent on the number of story readings,  $F(4, 140) = 22.17$ ,  $MSE = .06$ . When the stories had not been read, performance was unaffected by fact framing. Without story reading, performance in the correct and misleading conditions never differed significantly from the neutral baseline.

As in Experiment 1, prior reading of correct facts aided performance. For both easy and hard questions, performance was significantly higher following one or two story readings. The increase from one to two story readings did not reach significance for easy questions,  $t(35) = 1.54$ ,  $SEM = .05$ , but did for hard questions,  $t(35) = 2.94$ ,  $SEM = .07$ .

The novel part of Experiment 2 involved the misinformation items. Interestingly, having read misinformation sometimes reduced ability to correctly answer the questions. After having read the misinformation twice, subjects answered fewer easy questions correctly than baseline,  $t(35) = 3.11$ ,  $SEM = .07$ . Only in this cell was the reduction significant; performance for hard questions was approaching floor, making it difficult to obtain a significant reduction.

#### Production of misinformation

A 2 (question difficulty: easy or hard)  $\times$  3 (fact framing: correct, neutral, or misleading)  $\times$  3 (number of story readings: 0, 1, or 2) ANOVA model was conducted on proportion of questions answered with target misinformation. The data are shown in the top panel of Table 3. Misinformation was slightly more likely to be produced for hard questions than easy,  $F(1, 35) = 3.95$ ,  $MSE = .03$ . Misinformation was overwhelmingly produced for questions corresponding to facts that had had a misleading frame,  $F(2, 70) = 68.34$ ,  $MSE = .04$ , and following multiple story readings,  $F(2, 70) = 13.99$ ,  $MSE = .04$ . The only significant interaction was the critical one between fact framing and number of story readings,  $F(4, 140) = 32.73$ ,  $MSE = .02$ . For easy questions, production of misinformation increased above baseline both after both one,  $t(35) = 2.60$ ,  $SEM = .05$ , and two readings of misinformation,  $t(35) = 6.59$ ,  $SEM = .05$ . For hard questions, production of misinformation increased above baseline after both one,  $t(35) = 4.59$ ,  $SEM = .05$ , and two story readings,  $t(35) = 7.94$ ,  $SEM = .04$ . Misinformation production increased further increased from one to two story readings, for both easy,  $t(35) = 1.99$ ,  $SEM = .07$ ,  $p < .06$ , and hard questions,  $t(35) = 2.50$ ,  $SEM = .05$ .

#### Source judgments: Story attributions

To begin, we examined subjects' awareness of whether correct answers had been read in the stories. The middle panel of Table 2 shows the probability of a story attri-



Table 3

Mean proportion of questions answered with misinformation, as a function of question difficulty, number of story readings, and experimental condition

	Easy questions			Hard questions		
	Number of story readings			Number of story readings		
	Zero	One	Two	Zero	One	Two
Exp. 2						
Correct	.04	.06	.00	.10	.03	.05
Neutral	.05	.10	.05	.07	.06	.09
Misleading	.07	.22	.36	.06	.30	.43
Exp. 3						
Immediate						
Correct	.04	.01	.04	.08	.06	.09
Misleading	.06	.21	.35	.06	.26	.42
Delay, not tested						
Correct	.03	.04	.04	.07	.07	.05
Misleading	.06	.07	.13	.08	.10	.11
Delay, tested						
Correct	.05	.01	.04	.08	.06	.04
Misleading	.04	.13	.22	.09	.19	.25

bution, given production of the correct answer. These data are conditional probabilities, and thus not all subjects are included in each mean. Similar results were obtained when the joint probabilities were analyzed.

As in Experiment 1, subjects rarely made story attributions when they had not read the relevant stories. Story attributions increased dramatically when subjects had read the stories. Reading neutral frames and misinformation answers did lead to false claims of having read the correct answer in the story. However, these conditional probabilities were lower than those observed after having read correct answers in the stories. The conditional probability of a story attribution was greater after having read the correct answer to easy questions

once or twice than after having read the misinformation once,  $t(27) = 2.05$ ,  $SEM = .09$ , or twice,  $t(22) = 4.59$ ,  $SEM = .10$ . The conditional probability of a story attribution was greater after having read the correct answer to a hard questions twice than after reading the misinformation twice,  $t(8) = 2.92$ ,  $SEM = .15$ .

We next examined subjects' awareness of having read the misinformation answers in the stories. The top panel of Table 4 shows the probability of a story attribution, given production of the target misinformation answer. As for correct answers, subjects rarely made story attributions when they had not read the relevant stories. Prior reading correct and neutral frames did not lead to claims of having read misinformation in the stories. In sharp

Table 4

Proportion of misinformation answers that subjects remembered reading in the stories, as a function of question ease, number of story readings, and experiment condition

	Easy questions			Hard questions		
	Number of story readings			Number of story readings		
	Zero	One	Two	Zero	One	Two
Exp. 2						
Correct	.00	.14	.00	.08	.00	.00
Neutral	.00	.13	.00	.00	.00	.00
Misleading	.00	.52	.74	.00	.60	.97
Exp. 3						
Delay, not tested						
Correct	.00	.00	.00	.00	.00	.00
Misleading	.00	.00	.11	.00	.00	.53
Delay, tested						
Correct	.00	.04	.00	.00	.00	.05
Misleading	.00	.00	.50	.00	1.0	.54

contrast, story attributions increased dramatically when subjects had read misinformation in the stories. Subjects were aware of having read the misinformation answers in the stories. Because the rate of misinformation production was so infrequent except when subjects had read the misinformation, no statistics were computed on these conditional probabilities. Statistical significance was observed when the analyses were conducted on the joint probabilities,  $F(4, 140) = 38.19, MSE = .01$ .

*Source attributions: General knowledge attributions*

Of interest was whether subjects thought they were drawing on pre-experimental knowledge. A 2 (question difficulty: easy or hard)  $\times$  3 (fact framing: correct, neutral, or misleading)  $\times$  3 (number of story readings: 0, 1, or 2) ANOVA was conducted on proportion of answers

that were both correctly answered and attributed to general world knowledge. The relevant data are shown in Fig. 2. For present purposes, of main interest is the significant interaction between fact framing and number of story readings,  $F(4, 140) = 8.59, MSE = .07$ . As in Experiment 1, reading correct facts in the stories increased answers that were both correct and attributed to prior knowledge. Reading correct information twice increased later production of correct answers that were attributed to general world knowledge, for both easy,  $t(35) = 4.45, SEM = .05$ , and hard questions,  $t(35) = 2.45, SEM = .06$ .

Of particular interest were subjects' attributions of misinformation answers, since these incorrect answers were unlikely to have been part of people's general world knowledge prior to the experiment. Fig. 3 shows the

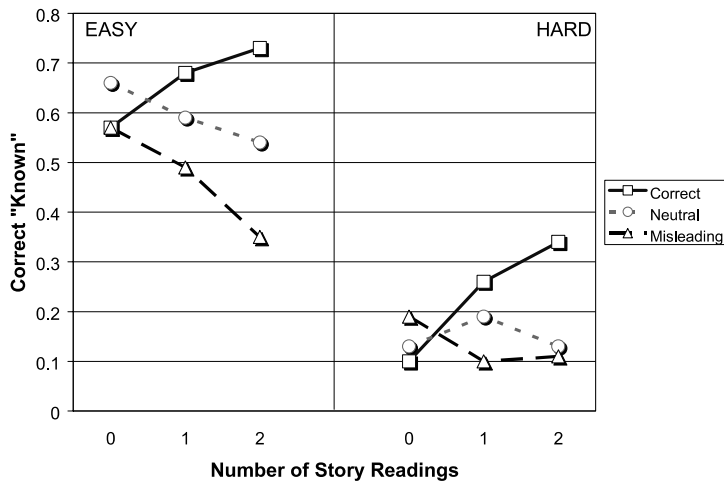


Fig. 2. Joint probability of answering a question correctly and saying 'yes' the answer was known prior to the experiment (Exp. 2).

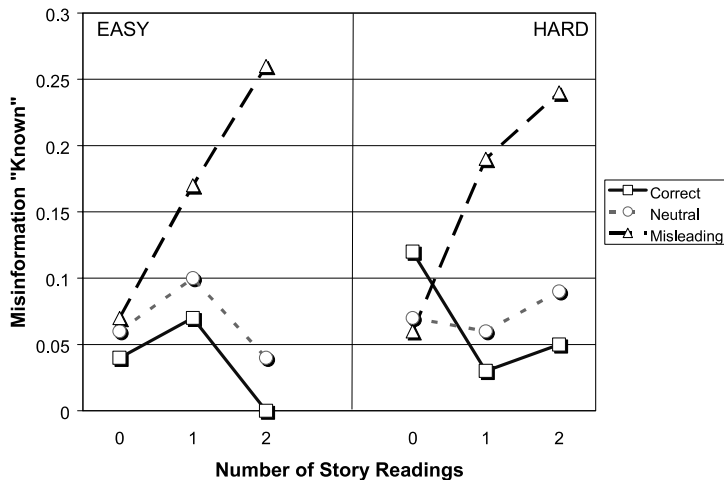


Fig. 3. Joint probability of answering a question with the target misinformation and saying 'yes' the answer was known prior to the experiment (Exp. 2).

joint probabilities of producing the target misinformation and saying ‘yes’ the answer was known prior to the experiment. Again of primary importance was the significant interaction between fact framing and number of story readings,  $F(4, 140) = 14.12$ ,  $MSE = .02$ . Only prior reading of misinformation led to a significant increase in misinformation attributed to general knowledge. This occurred for both easy,  $t(35) = 3.73$ ,  $SEM = .05$ , and hard questions,  $t(35) = 4.12$ ,  $SEM = .04$ . That is, subjects experienced an illusion of prior knowledge, claiming that they knew all along misinformation that was largely learned from the stories.

In summary, subjects were aware of their reliance on fictional sources. Critically, however, reading facts in fiction created an illusion of prior knowledge. Subjects thought they had known both correct and target misinformation answers all along.

#### Surprise ratings

A 2 (question difficulty: easy or hard)  $\times$  3 (number of story readings: 0, 1, or 2) ANOVA was conducted on mean surprise ratings. As expected, subjects were more surprised by the correct answers to difficult than easy questions,  $F(1, 35) = 52.02$ ,  $MSE = .15$ . Surprise was also dependent on the number of story readings,  $F(2, 70) = 9.42$ ,  $MSE = .16$ . The interaction was not significant,  $F(2, 70) = 1.68$ ,  $MSE = .14$ . For both easy and hard questions, subjects were more surprised by correct answers when they had read misinformation once or twice than when they had not read it.

#### Awareness of misinformation

Overall, almost two-thirds of subjects ( $n = 23$ ) reported some level of awareness of the misinformation. On average, subjects were able to specify less than one item on which they had been misled ( $M = .91$ ).

#### Discussion of Experiment 2

In Experiment 2, story reading led to the production of misinformation answers on the test of general world knowledge. Misinformation reduced correct answers below the neutral baseline, suggesting that in at least some instances, subjects were changing their answers to match the misinformation. Critically, however, source judgments suggested that while subjects were aware of their reliance on fictional sources, they also believed they had known these wrong answers prior to the experiment.

The story source data also allow us to make an interesting side point about false memories. While story attributions were always highest for correct and misinformation answers read in the stories, subjects also erroneously claimed to have read non-presented correct answers in the stories. That is, having read a neutral or misleading frame led to the erroneous belief that the

correct answer had been read in the story (see Table 2). A similar effect was not observed for misinformation answers (see Table 4). Possibly subjects generated the correct answers upon encountering a neutral or misleading frame in the story (e.g., spontaneously thought “London” when reading about Heathrow airport). The later misattribution of these inferences to the stories would yield the observed pattern of false alarms. The low base rate production of misinformation suggests subjects were unlikely to have spontaneously thought of the target misinformation when reading a correct or neutral frame, and hence misinformation answers were rarely incorrectly attributed to the stories. The results are similar to the false memories observed after reading a list of related words (Roediger & McDermott, 1995) or reading statements that only implied a particular fact (Brewer, 1977).

### Experiment 3

In Experiment 3, we examined the persistence of our effects a week after story reading. The effects of delay are of interest for two reasons. First, educational assessment often occurs long after learning. Second, people are more likely to confuse sources over time (e.g., Underwood & Pezdek, 1998). We wondered whether subjects would forget having read the answers in the stories after a week’s delay. Also of interest was what would insulate the effects against the passage of time. We chose to examine the effect of taking an initial test, a manipulation known to boost later performance (e.g., see Roediger & Gynn, 1996). Thus, Experiment 3 involved two sessions, separated by one week. In session one, subjects read the stories and took a general knowledge test that included questions on half the critical facts. In session two, subjects took a general knowledge test that included all of the critical questions and required source judgments.

#### Method

##### Subjects

Sixty Washington University undergraduates participated in the experiment for partial fulfillment of a course requirement. Twelve subjects were eliminated, either because they failed to finish reading at least 7 of the 9 stories, or because they did not return for the second session of the experiment. Thus, 48 subjects were included in the analyses.

##### Design

The experiment consisted of a 2 (question difficulty: easy or hard)  $\times$  2 (framing of facts: correct or misleading)  $\times$  2 (initial testing of facts: yes or no)  $\times$  3 (number of story readings: 0, 1, or 2) within-subjects design.

Of interest were subjects' responses on the test of general world knowledge at a 7-day delay, and the corresponding source attributions.

### Materials

The same 9 stories were used, with minor changes to the critical items. Each story contained four facts corresponding to easy questions and four facts corresponding to difficult questions. Half the frames were read with the correct answers and half with misleading answers.

Two new general knowledge tests were used in Session I. Each version of the initial test included 84 items. The test contained 48 filler questions of varying difficulty, plus the 36 critical questions. The 36 critical questions consisted of 2 easy (corresponding to 1 correct and 1 misleading fact) and 2 hard questions (corresponding to 1 correct and 1 misleading fact) from each of the 9 stories; there were two different session I tests in order to counterbalance which critical items were tested initially.

The general knowledge test used in Session II was the same as in Experiment 2, with the addition of a few questions corresponding to the increase in critical questions from Experiment 2. There were a total of 176 items on the test used in Session II.

The only other new material created for Experiment 3 was a brief questionnaire that assessed whether subjects had researched any of the facts over the delay.

### Procedure

Experiment 3 involved two experimental sessions one week apart. Session I closely followed the procedure of Experiment 2. Subjects read the stories, completed the filler task, and took the initial general knowledge test. Because of the shorter length of the test, subjects were given 10 min to complete the test, and no source judgments were made.

Seven days later, subjects returned to the lab. They first completed a brief math filler task (which was part of the reasoning experiment cover story). Subjects then took the final general knowledge test. They received 22 min in which to answer all of the questions. The source instructions were then read and subjects had 12 min to work through the test a second time making their source judgments. For each item, subjects made separate yes–no responses for the story and general knowledge sources.

As in Experiment 2, subjects did a number of tasks that assessed their awareness of the experiment's purpose. They answered the questionnaire assessing their knowledge of the misinformation, and the questionnaire probing their behavior during the delay. During debriefing, they rated how surprised they were by each of the corrected facts previously read in misleading form.

## Results

### Effects of delay

To assess whether the effects of reading fiction persisted over time, we compared performance on the immediate test to that on Session II questions that had *not* been tested previously. Thus, we computed 2 (time of testing: immediate or delayed)  $\times$  2 (question difficulty: easy or hard)  $\times$  2 (fact framing: correct or misleading)  $\times$  3 (number of story readings: 0, 1, or 2) ANOVAs on proportions of questions answered correctly versus with misinformation.

We begin with correct answers. All significant lower-order effects will be discussed in the context of the significant three-way interaction between test-time, fact framing, and number of story readings,  $F(2, 94) = 7.48$ ,  $MSE = .05$ . An examination of the bottom portion of Table 1 (the Immediate and Delayed Not Tested conditions) suggests that the effects of story exposure were greater on the immediate than the delayed test. To investigate this, separate ANOVAs were conducted on immediate vs. delayed (but not previously tested) items. The interaction between fact framing and number of story readings reached significance for the immediate test,  $F(2, 94) = 30.11$ ,  $MSE = .05$ , but was only marginally so for the delayed test,  $F(2, 94) = 2.40$ ,  $MSE = .05$ ,  $p < .1$ .

Replicating previous findings, story reading strongly affected performance on the immediate test. Reading correct answers significantly increased answers above baseline, for both easy and hard questions. Reading misinformation decreased ability to correctly answer easy questions. While effects on the immediate test were much stronger than those observed on the delayed test, some effects persisted over time. After a delay, subjects did answer more easy questions correctly after having read correct information once,  $t(47) = 2.15$ ,  $SEM = .05$ , or twice,  $t(47) = 2.05$ ,  $SEM = .04$ . They also showed some cost of misinformation, answering fewer easy questions correctly after having reading misinformation twice,  $t(47) = 2.08$ ,  $SEM = .06$ . For hard questions, no effects of story exposure remained after one week. Thus, delay reduced the effects of story reading on correct answers, although some effects (both positive and negative) remained for easy questions.

We turn now to the effect of delay on misinformation answers; the relevant data are in the bottom panel of Table 3 (the Immediate and Delayed Not Tested conditions). Overall, there was a main effect of test-time; more misinformation was produced on the immediate test than the delayed,  $F(1, 47) = 36.77$ ,  $MSE = .04$ . Test-time interacted with both fact framing and number of story readings; the three-way interaction was also significant,  $F(2, 94) = 21.10$ ,  $MSE = .02$ . To explore this result, separate ANOVAs were conducted on immediate vs. delayed items. The results from the immediate

condition replicated prior findings; subjects produced significantly more misinformation after having read it in the stories. In contrast, after a delay, without retrieval practice, story exposure did not significantly increase the production of misinformation items. Other than the main effects of fact framing and difficulty of question, nothing was significant.

In summary, we replicated strong effects of story reading on the immediate test, both in correct and target misinformation answers. After a delay, effects were greatly reduced. Effects were limited to easy questions; prior reading of correct answers boosted performance on these items whereas it was impaired by prior reading of misinformation.

#### *Insulating effects of prior exposure*

Did prior testing protect against the effects of time? To answer this question,  $2$  (prior testing: yes or no)  $\times 2$  (question difficulty: easy or hard)  $\times 2$  (fact framing: correct or misleading)  $\times 3$  (story reading: 0, 1, or 2) ANOVAs were conducted on answers on the Session II test.

We begin with correct answers. As expected, there were main effects of question difficulty and fact framing, which will be discussed in the context of higher-order interactions. The relevant data are in the bottom panel of Table 1 (the two delayed conditions). The critical interaction between fact framing and number of story readings was significant,  $F(2, 94) = 11.89$ ,  $MSE = .06$ , and was qualified by a marginally significant interaction with prior testing,  $F(2, 94) = 2.59$ ,  $MSE = .06$ ,  $p < .1$ . As described earlier, if subjects were only tested at time II, the effects of story exposure (on correct answers) were limited to easy questions. A different pattern was obtained at a delay for items that also had been tested in the initial session. For easy questions, subjects answered more questions correctly if they had read the correct information once,  $t(47) = 2.54$ ,  $SEM = .05$ , or twice,  $t(47) = 3.09$ ,  $SEM = .06$ . As expected, the opposite effect occurred following misinformation. Subjects answered more questions correctly if they had not read the misinformation; performance dropped after reading the misinformation once,  $t(47) = 2.42$ ,  $SEM = .05$ , or twice,  $t(47) = 2.76$ ,  $SEM = .06$ . For hard questions, subjects benefited from reading the correct information once,  $t(47) = 3.43$ ,  $SEM = .05$ , or twice,  $t(47) = 3.09$ ,  $SEM = .06$ . However, there was no cost to the misinformation (on correct answers) for hard questions.

We turn now to the effect of prior testing on delayed production of misinformation. The relevant data are in the bottom panel of Table 3 (the two delayed conditions). There were the expected main effects of question difficulty and fact framing. Of interest was a significant interaction between prior testing and fact framing; prior testing served to increase production of misinformation,  $F(1, 47) = 12.31$ ,  $MSE = .02$ . The interaction between

fact framing and number of story readings was significant, and further qualified by a three-way interaction with prior testing,  $F(2, 94) = 5.65$ ,  $MSE = .02$ . As described earlier, one or two exposures to misinformation did not lead to its production after a delay (without prior testing). However, this effect changed when subjects had been tested previously on the items. When subjects had been previously tested on the relevant facts, misinformation effects were obtained following a one-week delay. For easy questions, misinformation production was boosted above baseline after one,  $t(47) = 1.95$ ,  $SEM = .02$ , or two readings,  $t(47) = 1.94$ ,  $SEM = .02$ . Misinformation was also produced for hard items. Subjects produced misinformation above baseline following one,  $t(47) = 2.59$ ,  $SEM = .03$ , or two readings,  $t(47) = 4.87$ ,  $SEM = .04$ . Misinformation was produced after a delay, but only for items that had been tested previously.

In summary, prior testing reduced the effects of delay. Even after a delay of one week, story reading affected both correct and misinformation answers—but only for those questions that were also tested in the first session. That is not to say, however, that prior testing preserved the effects at levels equivalent to those observed on the immediate test. Direct comparisons between immediate and delayed-previously tested items did suggest differences. That is, testing condition did interact significantly with fact framing and number of story readings, for both correct,  $F(2, 94) = 3.41$ ,  $MSE = .02$ , and misinformation answers,  $F(2, 94) = 8.40$ ,  $MSE = .01$ . For correct answers, an examination of Table 1 suggests two major differences. Easy questions showed a greater cost from two exposures to misinformation on the immediate than the delayed test,  $t(47) = 3.85$ ,  $SEM = .03$ . Hard questions showed a greater benefit from two exposures to correct information on the immediate than the delayed test,  $t(47) = 3.07$ ,  $SEM = .04$ . Turning to the misinformation, Table 3 suggests the differences between the immediate and delayed-prior-tested tests were ones of magnitude of errors, not their patterns. That is, both conditions yielded larger misinformation effects after reading misinformation once or twice; however, these effects were bigger in the immediate condition. After reading misinformation once, more misinformation was produced on the immediate than the delayed test for previously tested items, for both easy,  $t(47) = 2.7$ ,  $SEM = .03$ , and hard questions,  $t(47) = 2.28$ ,  $SEM = .03$ . Similarly, after reading misinformation twice, more misinformation was produced on the immediate than the delayed test for previously tested items, for both easy,  $t(47) = 3.51$ ,  $SEM = .04$ , and hard questions,  $t(47) = 3.93$ ,  $SEM = .04$ .

Thus, for both correct and misinformation answers, the largest effects of story reading occurred on the immediate test, and smaller but still significant effects

occurred for items that had been tested previously. Without prior testing, however, after a week's delay subjects no longer produced misinformation, and only showed benefits of story reading for easy questions.

#### Source attributions

Source judgments were only collected on the second, delayed test. We begin by discussing how aware subjects were of the story source, one week after story reading. The relevant data for correct answers are in the bottom panel of Table 2. When the relevant stories had not been read, story attributions were rare and did not differ across conditions. First, consider the pattern of data when subjects had *not* taken an initial test. Although numerically story attributions were always greater when subjects had read correct answers once or twice, this difference only reached significance in one case. When subjects had not been tested initially, they made more story attributions after twice reading correct answers to easy questions,  $t(37) = 3.30$ ,  $SEM = .06$ . In contrast, the patterns were much stronger when subjects had taken an initial test, as was the overall level of source awareness. Subjects who had been tested initially made more story attributions after reading consistent answers to easy questions once,  $t(39) = 3.70$ ,  $SEM = .07$ , or twice,  $t(41) = 5.08$ ,  $SEM = .07$ . A similar pattern was obtained when answers to hard questions had been read once,  $t(17) = 4.18$ ,  $SEM = .11$ , or twice,  $t(16) = 2.34$ ,  $SEM = .14$ . Thus, in summary, subjects still showed awareness of the story source after a delay, although this awareness was much greater if the subjects had been tested on the same questions during the first session. This same pattern was obtained (and statistical significance obtained) if the dependent measure involved the joint probabilities.

Subjects also showed some awareness of having read misinformation answers in the stories, albeit at lower levels than on previous immediate tests (Experiment 2). Table 4 shows the probability of making a story attribution, conditional upon production of the target misinformation answer. Each mean thus includes data only from those subjects who produced misinformation in that cell. However, the pattern of the data is clear: story attributions occurred only when subjects had actually read the misinformation in the stories. A similar pattern was observed (and backed by ANOVA statistics) when the dependent measure was the joint probability of producing misinformation and attributing it to the stories (as opposed to conditional probabilities).

Delay reduced but did not eliminate beliefs of prior knowledge. In particular, the illusion of prior knowledge was weak for correct answers; the relevant data are shown in Fig. 4. In most cases, rates of saying 'general knowledge' were not above baseline. Only in two cases was there an increase over baseline in correct answers attributed to general knowledge. Reading correct

information once increased correct answers attributed to prior knowledge, for both easy,  $t(47) = 1.88$ ,  $SEM = .06$ ,  $p < .07$ , and hard questions,  $t(47) = 2.04$ ,  $SEM = .05$ .

Of particular interest were prior knowledge attributions for misinformation answers, since these were rarely produced without story reading. A 2 (question difficulty: easy or hard)  $\times$  2 (fact framing: correct or misleading)  $\times$  3 (number of story readings: 0, 1, or 2) ANOVA was computed on general knowledge attributions (the sum of 'general knowledge and story' plus 'general knowledge' only attributions). This analysis was done only on items tested previously (since only for these was the misinformation effect significant), although all data are shown in Fig. 5. The critical interaction between fact framing and number of story readings was significant,  $F(2, 94) = 4.64$ ,  $MSE = .02$ . Only after reading misinformation did subjects produce more misinformation that was attributed to general world knowledge. For easy questions, there was a trend for reading misinformation twice to increase general knowledge estimates above baseline,  $t(47) = 1.83$ ,  $SEM = .03$ . Similar results were obtained for hard questions,  $t(47) = 1.74$ ,  $SEM = .03$ . Reading misinformation inappropriately increased general knowledge attributions.

#### Surprise ratings

A 2 (question difficulty: easy or hard)  $\times$  3 (number of story readings: 0, 1, or 2) ANOVA was conducted on mean surprise ratings. Subjects again rated the answers to easy questions as less surprising than answers to hard questions,  $F(1, 47) = 113.86$ ,  $MSE = .17$ . Surprise increased with number of story readings,  $F(2, 94) = 7.15$ ,  $MSE = .10$ , but the interaction between story readings and question difficulty was not significant,  $F(2, 94) = 2.02$ ,  $MSE = .06$ . For both easy and hard questions, subjects were more surprised when they had read misinformation once or twice than when they had not read the relevant stories. Subjects apparently believed the misinformation, since they were more surprised by the correct answers to questions for which they had read misinformation.

#### Awareness of experiment

A week after story reading, 81% of subjects reported awareness that story information could be used to answer the general world knowledge questions. Seventy-five percent remembered reading misinformation in the stories, although they were only able to report (on average) 1.2 specific errors.

#### Activities during delay

Six subjects reported researching one question during the delay. Of these, only three were from the critical set. Thus, we are not concerned that subjects learned the answers in the interim between the two tests.

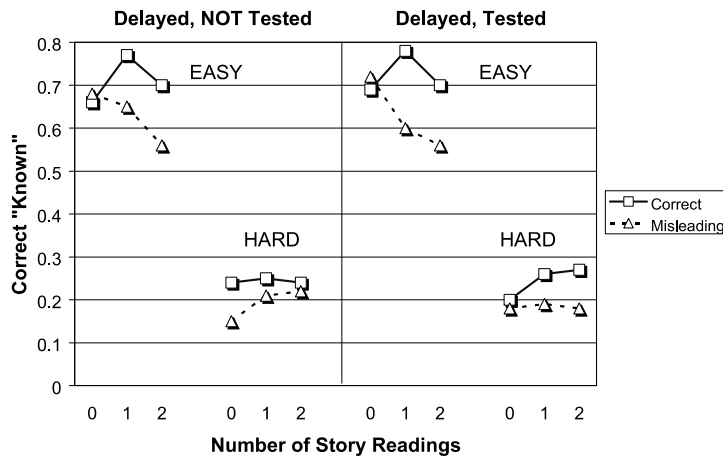


Fig. 4. Joint probability of answering a question correctly and saying 'yes' the answer was known prior to the experiment (Exp. 3).

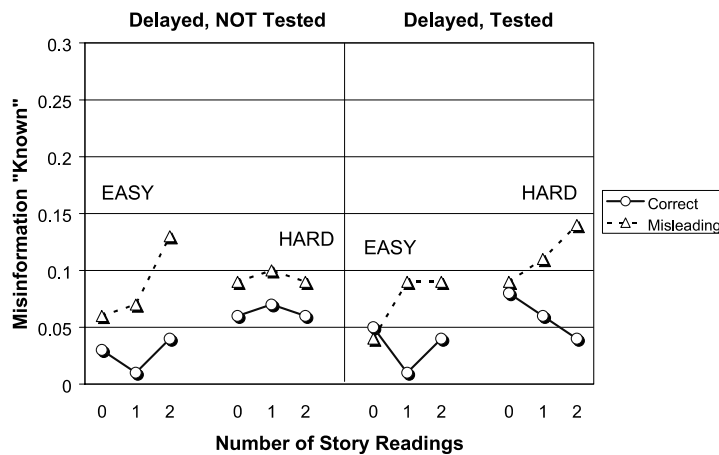


Fig. 5. Joint probability of answering a question with the target misinformation and saying 'yes' the answer was known prior to the experiment (Exp. 3).

### Discussion of Experiment 3

Replicating earlier results, story reading had a large impact on subjects' correct and target misinformation answers on the immediate test. A one-week delay clearly diluted the effects of story reading. However, some of the effects persisted over the delay; subjects answered more easy questions correctly if they had read the relevant stories, and answered fewer easy questions correctly after reading misinformation. They did not show the parallel effects for hard questions. Prior testing buffered effects against the passage of time and led to performance on easy and hard questions intermediate between that observed in the immediate and delayed (no prior testing) conditions.

The data from the two delayed conditions stress the importance of linking the answer in memory to the

question that will later serve at its retrieval cue. Easy questions, answers, and lures are most likely to already be established in memory (although not necessarily already associated to each other). Reading the correct answer in the story creates or strengthens a link between two familiar concepts, leading to a boost in performance. Reading an incorrect answer creates or strengthens a different link, without necessarily destroying the pre-existing link between the cue and correct answer. On the other hand, difficult questions, answers, and lures are by definition lower frequency items, and thus less likely to be already established in memory. Even if the difficult question and (correct or incorrect) answer are read in association, this may not be sufficient to create a strong enough link to persist over delay. Retrieval practice (via immediate testing) strengthens these links, allowing persistence of effects

over delay. Retrieval practice allows subjects to establish alternatives to difficult questions, and to strengthen links to wrong answers for easy questions. Thus, the misinformation effects reflect not only suggestibility, but also the establishment in memory of question-answer pairings. Manipulations that help the subject to remember the misinformation (such as testing) increase suggestibility over a delay.

After a week's delay, we failed to eliminate the link to the story source. In Experiment 3, subjects still showed some knowledge of the story link, although clearly less than on immediate tests. Prior testing minimized but did not eliminate the effects of delay. Misinformation responses, however, were still misattributed to prior knowledge for items that had also been tested in the first session. Finally, note that the delayed source errors were similar to those observed in Experiment 2; correct answers produced after reading misinformation were misattributed to the stories, but misinformation answers produced after reading correct answers were not misattributed to the stories.

### General discussion

In three experiments, subjects used information from fictional stories to help them answer general knowledge questions. In all experiments, subjects correctly answered more questions when the answers had been read in the stories. In Experiments 2 and 3, subjects' behavior also showed the costs of misinformation. After reading misinformation, they answered fewer questions correctly and produced more target misinformation answers. Subjects were very aware that many of their answers had appeared in the stories. However, story reading also increased belief of prior knowledge. Critically, this effect also occurred for misinformation, which was unlikely to have been known prior to the experiment. With immediate testing, the positive and negative effects of story exposure were impressively large, even though subjects were cautioned against guessing on the general knowledge test. The effects were smaller after a week's delay, but testing at the immediate session helped minimize the effects of delay. Subjects believed their incorrect (suggested) answers, giving higher ratings of 'surprise' to items for which they had read misinformation than those from unread stories.

Overall, the data support the hybrid view of fact representation. That is, facts learned from fiction were linked in memory both to related pre-experimental knowledge and to the story source. Integration with other related knowledge is suggested by subjects' willingness to use story facts on a general knowledge test. In fact, they often believed they knew the story facts prior to the experiment (even for misinformation items, which should not have been 'known' prior to the experiment).

However, subjects were also very good at identifying which answers had come from the stories. Even after a delay, facts learned from fiction still contained strong links to the story sources.

Having read an answer in the stories probably increased the ease with which it came to mind at test, leading to its production and confidence in its correctness (e.g., Kelley & Lindsay, 1993). As exemplified in the effects of misinformation, story reading also led to the creation and strengthening of new associations, which then overrode prior associations. After reading misinformation, the ability to answer easy questions dropped below the neutral baseline. That is, exposure to misinformation reduced subjects' ability to answer easy questions that they otherwise should have been able to answer. A new question-answer pairing was stronger than the old.

The hybrid representation hypothesis helps us to understand why subjects relied on the stories even though they often noticed that the stories contained errors. It was not necessary for subjects to forget the source of their knowledge. Rather, what was critical was that story reading led to an illusion of truth. Misinformation production was very low prior to the experiment and thus should not have been attributed to general knowledge. However, reading misinformation led to its production, and these answers were often attributed to pre-experimental knowledge. This illusion of prior knowledge is similar to the knew-it-all-along bias (Fischoff, 1977; Wood, 1978) exhibited by people who have been told correct answers or otherwise received feedback about judged events. With the answer in front of them, people overestimate how likely they would have been to produce the answer. In this case, once subjects produced the misinformation, they were unable to judge their prior knowledge.

Contrary to expectations, we found little evidence for monitoring behavior. Story reliance did not decrease when misinformation was added to the stories. An informal cross-experiment comparison reveals benefits from having read correct information in Experiment 2 for both easy (+.27 over the neutral baseline) and hard questions (+.43) even though each story contained multiple errors. The presence of misinformation did not make subjects less likely to pull correct answers from the story, even though the post-experiment questionnaire indicated that subjects were aware that the stories contained some errors. And subjects were not selective in their use of story information—strong misinformation effects occurred for easy as well as hard questions. This is surprising since obviously incorrect information could have been expected to serve as warnings to subjects that the stories were unreliable (e.g., Loftus, 1979). Even if prior reading eased retrieval of misinformation at test, we had expected subjects to monitor the source of this fluency and effectively exclude it. Normally, young



adults are able to use episodic memory to assess the source of fluency, *unless* they studied under divided attention conditions (e.g., Jacoby et al., 1998; Kelley & Lindsay, 1993, exp. 4a). In contrast, our subjects showed increased reliance on repeated story facts without the burden of divided attention, presumably because repetition also increased the illusion of prior knowledge.

We can only speculate about why Gerrig and Prentice (1991) found evidence for monitoring and selective use of story information in their studies while we did not. When reading for pleasure, one may be less likely to engage in the critical processing necessary to notice, respond to, and reject the misinformation (Prentice & Gerrig, 1999). Gilbert has argued that the default mode is to believe information, and that to ‘unbelieve’ information takes effort (Gilbert, 1991; Gilbert, Krull, & Malone, 1990). Indeed, readers deeply involved in text are less likely to detect false notes (Green & Brock, 2000). If our subjects did not notice the errors in the stories, these errors would come to mind fluently at test without an accompanying ‘warning’ of falsehood. However, this line of reasoning suggests our subjects were more involved in our stories than were Gerrig’s in his stories, and we have no evidence (or reason) to support that assumption. Figuring out when subjects do vs. do not monitor (and selectively use) fiction remains a question for further research.

We close by noting the relationship of our results to those from other related experimental paradigms (see also Marsh & Bower, 1999). As we have noted repeatedly, our procedures parallel the eyewitness post-event information paradigm (e.g., Loftus et al., 1978; see also Meade & Roediger, 2002; Roediger, Meade, & Bergman, 2001), but the results are not the same—in the Loftus paradigm, a spillover effect often occurs once subjects detect misinformation, reducing suggestibility. In our paradigm, no such spillover effect occurred, even though a large amount of misinformation was included across stories. It is also not clear that the underlying mechanisms are the same in the two paradigms. In the eyewitness paradigm, drawing subjects’ attention to source information reduces suggestibility (e.g., Lindsay & Johnson, 1989), whereas it does not in our paradigm. Subjects in the eyewitness paradigm may be more likely to try and remember the specific study phase, whereas subjects answering general knowledge questions may answer based on what comes easily to mind.

Our data also parallel illusory truth effects (Hasher, Goldstein, & Toppino, 1977), in which the mere presentation of a fact increases its rated truth. Again, however, the results are not the same—in the illusory truth effect, subjects are able to take advantage of source information to reduce the effect (e.g., Brown & Nix, 1996). Fiction is a rather unusual source—knowing that an item came from a fictional source does not necessarily mean that it is wrong. In addition, as suggested earlier,

fiction may be a special case in that subjects approach it less critically than other sources. Thus, large effects occur from fiction reading, even in circumstances that would reduce suggestibility in other paradigms.

## References

- Begg, I. M., Robertson, R. K., Gruppasop, V., Anas, A., & Needham, D. R. (1996). The illusory-knowledge effect. *Journal of Memory and Language*, *35*, 410–433.
- Blaxton, T. A. (1989). Investigating dissociations among memory measures: Support for a transfer-appropriate processing framework. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*, 657–668.
- Brewer, W. F. (1977). Memory for the pragmatic implications of sentences. *Memory & Cognition*, *5*, 673–678.
- Brown, A. S., & Nix, L. A. (1996). Turning lies into truths: Referential validation of falsehoods. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*, 1088–1100.
- Conway, M. A., Gardiner, J. M., Perfect, T. J., Anderson, S. J., & Cohen, G. M. (1997). Changes in memory awareness during learning: The acquisition of knowledge by Psychology undergraduates. *Journal of Experimental Psychology: General*, *126*, 393–413.
- Dodd, D. H., & Bradshaw, J. M. (1980). Leading questions and memory: Pragmatic constraints. *Journal of Verbal Learning and Verbal Behavior*, *19*, 695–704.
- Dubeck, L. W., Bruce, M. H., Schmuckler, J. S., Moshier, S. E., & Boss, J. E. (1990). Science fiction aids science teaching. *The Physics Teacher*, 316–318.
- Fischhoff, B. (1977). Perceived informativeness of facts. *Journal of Experimental Psychology: Human Perception and Performance*, *3*, 349–358.
- Gerrig, R. J., & Prentice, D. A. (1991). The representation of fictional information. *Psychological Science*, *2*, 336–340.
- Gilbert, D. T. (1991). How mental systems believe. *American Psychologist*, *46*, 107–119.
- Gilbert, D. T., Krull, D. S., & Malone, P. S. (1990). Unbelieving the unbelievable: Some problems in the rejection of false information. *Journal of Personality and Social Psychology*, *59*, 601–613.
- Green, M. C., & Brock, T. C. (2000). The role of transportation in the persuasiveness of public narratives. *Journal of Personality and Social Psychology*, *79*, 701–721.
- Hasher, L., Goldstein, D., & Toppino, T. (1977). Frequency and the conference of referential validity. *Journal of Verbal Learning and Verbal Behavior*, *16*, 107–112.
- Hoffman, H. G., Granhag, P. A., See, S. T. K., & Loftus, E. F. (2001). Social influences on reality-monitoring decisions. *Memory & Cognition*, *29*, 394–404.
- Hovland, C. I., Lumsdaine, A. A., & Sheffield, F. D. (1949). *Experiments on mass communication*. Princeton: Princeton University Press.
- Jacoby, L. L., Jones, T. C., & Dolan, P. O. (1998). Two effects of repetition: Support for a dual-process model of knowledge judgments and exclusion errors. *Psychonomic Bulletin & Review*, *5*, 705–709.
- Johnson, M. K., Raye, C. L., & Durso, F. T. (1980). Reality monitoring: Second perceptions and thoughts. *Bulletin of the Psychonomic Society*, *15*, 402–404.

- 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.
- Lewis, C. H., & Anderson, J. R. (1976). Interference with real world knowledge. *Cognitive Psychology*, 8, 311–335.
- Lindsay, D. S., & Johnson, M. K. (1989). The eyewitness suggestibility effect and memory for source. *Memory & Cognition*, 17, 349–358.
- Loftus, E. F. (1979). Reactions to blatantly contradictory information. *Memory & Cognition*, 7, 368–374.
- Loftus, E. F., Miller, D. G., & Burns, H. J. (1978). Semantic integration of verbal information into a visual memory. *Journal of Experimental Psychology: Human Learning and Memory*, 4, 19–31.
- Marsh, E. J., & Bower, G. H. (1999). Applied aspects of source monitoring. *Cognitive Technology*, 4, 4–17.
- Meade, M. L., & Roediger III, H. L. (2002). Explorations in the social contagion of memory. *Memory & Cognition*, 30, 995–1009.
- Nelson, T. O., & Narens, L. (1980). Norms of 300 general-information questions: Accuracy of recall, latency of recall, and feeling-of-knowing ratings. *Journal of Verbal Learning and Verbal Behavior*, 19, 338–368.
- Peterson, S. B., & Potts, G. R. (1982). Global and specific components of information integration. *Journal of Verbal Learning and Verbal Behavior*, 21, 403–420.
- Potts, G. R., & Peterson, S. B. (1985). Incorporation versus compartmentalization in memory for discourse. *Journal of Memory and Language*, 24, 107–118.
- Potts, G. R., St. John, M. F., & Kirson, D. (1989). Incorporating new information into existing world knowledge. *Cognitive Psychology*, 21, 303–333.
- Prentice, D. A., & Gerrig, R. J. (1999). Exploring the boundary between fiction and reality. In S. Chaiken, & Y. Trope (Eds.), *Dual process theories in social psychology* (pp. 529–546). NY: Guilford Press.
- 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.
- Roediger III, H. L., & Guynn, M. J. (1996). Retrieval processes. In E. L. Bjork, & R. A. Bjork (Eds.), *Memory* (pp. 197–236). San Diego: Academic Press.
- Roediger III, H. L., & McDermott, K. B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21, 803–814.
- Roediger III, H. L., Meade, M. L., & Bergman, E. T. (2001). Social contagion of memory. *Psychonomic Bulletin & Review*, 8, 365–371.
- Smith, J. A. (1993). Content learning: A third reason for using literature in teaching reading. *Reading, Research, and Instruction*, 32, 64–71.
- Smith, V. L., & Ellsworth, P. C. (1987). The social psychology of eyewitness accuracy: Misleading questions and communicator expertise. *Journal of Applied Psychology*, 72, 294–300.
- Storey, D. C. (1982). Reading in the content areas: Fictionalized biographies and diaries for social scientists. *Reading Teacher*, 35, 796–798.
- Strange, J. J., & Leung, C. C. (1999). How anecdotal accounts in news and fiction can influence judgments of a social problem's urgency, causes, and cures. *Personality and Social Psychology Bulletin*, 25, 436–449.
- Thapar, A., & Rouder, J. N. (2001). Bias in conceptual priming. *Psychonomic Bulletin & Review*, 8, 791–797.
- Underwood, J., & Pezdek, K. (1998). Memory suggestibility as an example of the sleeper effect. *Psychonomic Bulletin & Review*, 5, 449–453.
- Wheeler, S. C., Green, M. C., & Brock, T. C. (1999). Fictional narratives change beliefs: Replications of Prentice, Gerrig, and Bailis (1997) with mixed corroboration. *Psychonomic Bulletin & Review*, 6, 136–141.
- Wood, G. (1978). The knew-it-all-along effect. *Journal of Experimental Psychology: Human Perception and Performance*, 4, 345–353.