JOURNAL OF MEMORY AND LANGUAGE **35,** 300–318 (1996)
ARTICLE NO. 0017

Misinformation Effects in Recall: Creating False Memories through Repeated Retrieval

HENRY L. ROEDIGER III, J. DEREK JACOBY, AND KATHLEEN B. MCDERMOTT

Rice University

In two experiments subjects viewed slides depicting a crime and then received a narrative containing misleading information about some items in the slides. Recall instructions were manipulated on a first test to vary the probability that subjects would produce details from the narrative that conflicted with details from the slides. Two days later subjects returned and took a second cued recall test on which they were instructed to respond only if they were sure they had seen the item in the slide sequence. Our interest was in examining subjects' production of the misleading postevent information on the second cued recall test (on which they were instructed to ignore the postevent information) as a function of instructions given before the first test. In both experiments, robust misinformation effects occurred, with misrecall being greatest under conditions in which subjects had produced the wrong detail from the narrative on the first test. In this condition subjects were more likely to recall the wrong detail on the second test and were also more likely to say that they remembered its occurrence, when instructed to use Tulving's (1985) *remember*/*know* procedure, than in comparison conditions. We conclude that a substantial misinformation effect occurs in recall and that repeated testing increases the effect. False memories may arise through repeated retrieval.

The memory illusion that has received the most attention over the past 20 years is the deleterious effect of presenting people with misleading information after they have witnessed an event (e.g., Loftus, 1979a, 1979b). An experiment by Loftus, Miller, and Burns (1978) is typical. Subjects saw a series of 30 colored slides depicting an automobile accident in which a car failed to yield the right of way and a collision occurred. A critical detail in one of the slides was either a stop sign or a yield sign. (Across subjects each sign occurred equally often.) Subjects were then asked a series of questions about the slides. Some subjects were not asked about the sign (the control condition), whereas two other groups of subjects were asked ''Did the car pass the red Datsun while it was stopped at the stop sign?'' (or ''at the yield sign'' for the second group). These questions provided information that was either consistent with the detail in the original event or inconsistent (misleading) with respect to that detail. After a filler task, subjects were given a two-alternative forced choice recognition test, on which one question addressed the type of sign appearing at the intersection (i.e., stop or yield) in the original slide sequence. When no prior question about the sign had been asked, 63% of the subjects answered the question correctly, but when the prior question sequence had contained information consistent with the slides, correct recognition grew to 70%. Of most interest was the case in which inconsistent or misleading information had been given during the initial questions; accurate recognition was only 43% in this case. The misleading information given during the questions interfered with recognition of the correct answer. The Loftus et al. (1978) interpretation

This research was supported by Grant F49620-92-J-0437 awarded by the Air Force Office of Scientific Research to Henry L. Roediger III. We appreciate Maria Zaragoza providing us with the materials used in this research. We thank Doug L. Nelson and Maria Zaragoza for comments on a previous version of the manuscript. Reprint requests should be addressed to Henry L. Roediger III at the Department of Psychology, MS 25, Rice University, 6100 Main Street, Houston, TX 77005-1892. E-mail: roddy@rice.edu.

0749-596X/96 $18.00

was that the memory of this feature from the slides was altered, overwritten, or replaced by the misleading information.

The influence of misleading information on recognition and recall has been consistently obtained in many experiments, although the magnitude of the effect depends on many features of the experimental design and the type of test (see Loftus, 1979a, 1993 for reviews). The interpretation of the influence as due to impairment of the original memory, or overwriting, has been questioned by further experiments (McCloskey & Zaragoza, 1985), but the power of misleading information to alter retrieval is not in question. Furthermore, as Johnson, Hashtroudi, and Lindsay (1993) noted, although the fate of the original memory, or the memory impairment issue, is interesting, an examination of when and how people come to remember the suggested detail as having occurred in the original event is equally interesting. Johnson et al. (1993) argue that the source monitoring framework provides a useful approach to explaining and predicting misinformation effects. Briefly, when presented with misleading postevent information, subjects may later become confused as to the source of their memories (e.g., whether they saw a stop sign in the slide sequence or just read about it in the later questions). Such confusions can lead to subjects' misattributing a memory from one source as having occurred in another.

The current experiments examine a rather neglected feature of the eyewitness memory paradigm, viz., the influence of repeated testing on the probability with which subjects report having seen the suggested detail in the original event. The neglect of this variable seems surprising because as Poole and White (1995) have pointed out, witnesses to crimes are typically queried many times before they are asked to testify in court. If the act of repeated retrieval has systematic effects on memory in the eyewitness situation, the implications would be important for both theoretical and applied reasons. Poole and White (1995) summarized results of many experiments with children showing that repeated interviews can have both facilitating and detrimental effects on later retention. On the positive side, interviewing a child soon after he or she witnesses an event can have beneficial effects on later recall (e.g., Dent & Stephenson, 1979). However, children sometimes recall more erroneous information over repeated interviews, too (e.g., Poole & White, 1991). Before turning to the rationale for the current experiments, we first briefly summarize the effects of testing on retention in other paradigms with adult subjects.

The act of testing memory not only measures retention, but irrevocably alters future retention for the event that was tested. This fact has been demonstrated many times, and perhaps the most commonly studied manifestation is referred to as the *testing effect:* the act of recalling or recognizing material generally increases the likelihood of its later recall or recognition, relative to a control condition in which material is tested only at the later time (e.g., Spitzer, 1939; see also Hogan & Kintsch, 1970; Thompson, Wenger, & Bartling, 1978; Wheeler & Roediger, 1992). In a somewhat different research tradition, repeated testing of material not only prevents its forgetting, but also leads to greater overall recall on each test, a phenomenon that Erdelyi and Becker (1974) labeled hypermnesia (see Payne, 1987, for a review). This outcome has been demonstrated in an eyewitness memory paradigm, too, under conditions in which no misinformation was given (Scrivener & Safer, 1988).

Paradoxically, the act of testing can also have detrimental effects on later recall and recognition of material, and again several different lines of research establish this general conclusion. The phenomenon of output interference refers to the fact that the act of recalling some information inhibits later recall of other information, both in short-term memory situations (Tulving & Arbuckle, 1963) and in those testing long-term memory (e.g., Smith, 1971; Roediger & Schmidt, 1980). These and other inhibitory effects of recall have been reviewed in different eras by Roediger (1974),

Roediger and Neely (1982), and Anderson and Neely (in press).

The interest in the current experiments is whether there might be a testing effect in the basic eyewitness memory paradigm for misleading postevent information. If subjects are exposed to an original event (a slide sequence), and then receive misleading information with respect to that event, will they perform differently on a later recall test if there was an intervening test than if there was no intervening test? In addition, if subjects are induced to report an event that is in error on a first occasion, will this report alter performance on a later test (relative to a condition in which no inducing manipulation was introduced)? In short, there may be a testing effect for erroneously recalled information—production of an error may make it more likely to reappear on a later test. Furthermore, we may be able to manipulate the magnitude of the testing effect by altering response criterion on an initial test.

Evidence for a testing effect for false memories was reported by Roediger and McDermott (1995, Experiment 2), who gave subjects 16 lists of related words, each one derived from a common associate that was not presented. After 8 lists subjects immediately recalled as many words as they could, whereas after 8 others they did math problems. In the 8 recalled lists, the probability of recall of studied words was .62, and the probability of false recall of the critical nonpresented words (from which the lists were derived) was .55. Later subjects took a recognition test representing all 16 lists. The hit rate for studied items was .79 for lists that had been tested and .65 for those that had not; the false alarm rate for the critical items was .81 for lists that had been tested and .72 for lists that had not been tested. Both differences were reliable. Therefore, the testing effect occurred for both studied and nonstudied items, or for both veridical and false memories.[1]

Although Roediger and McDermott (1995) found positive effects of both accurate and inaccurate recall in boosting later hit and false alarm rates in recognition, others have not replicated this pattern. Payne, Elie, Neuschatz and Blackwell (1996) obtained a similar pattern for studied items, but—and critically for the present argument—they did not obtain an increase in recognition of the critical nonpresented items from prior testing on their delayed recognition test. Schacter, Pradere and Verfaellie (1996) obtained no effect of recall on later recognition for either the studied or the nonstudied items in the Roediger and McDermott (1995) paradigm.

The inconsistent results on these recognition tests may indicate that testing effects for false memories are slight or variable. This pattern would not be too surprising because some ambiguity exists as to whether a testing effect occurs in recognition for studied items after an initial recall test. The consensus is that there is at best a small testing effect on recognition after a single recall test. Darley and Murdock (1971) found little or no effect of prior free recall on later recognition of unrelated lists of words. Lockhart (1975) and Jones and Roediger (1995) did find an effect, but it was confined to the end of the list.

Although the effects of recall on later recognition of studied items seem variable, the effects of recall on a later recall test are almost always positive (see Roediger & Guynn, in press, for a partial review). The act of recalling material makes that material more recallable in the future, although the amount of benefit depends on the types of cues used on the two occasions (McDaniel, Kowitz, & Dunay, 1989). However, when a free recall test occurs on both occasions, prior recall always enhances later recall, relative to a condition in which there is no first test. Darley and Murdock (1971), who reported no

[1] When item analyses were conducted on the lists that had been immediately recalled, the finding was that if items (studied or critical) had been produced on the initial recall test, they were almost always recognized later (.98 for studied items and .93 for critical nonstudied items). For items that had not been recalled on the initial tests, the later hit rate and false alarm rates on the recognition test were .50 and .65, respectively. However, item selection difficulties hamper interpretation of these results.

effect of prior free recall on later recognition, did find a large effect of free recall on later free recall. McDermott (1996) had subjects study the lists developed by Roediger and McDermott (1995), taking a recall test on some lists and no test on the rest. She replicated the earlier findings of substantial false recall of the critical nonpresented words on the initial tests. Two days later subjects were tested by a final free recall test over all the studied lists. Taking a prior test had great effects on both accurate recall and false recall. Probability of false recall for the previously tested lists was .24, whereas false recall for the nontested lists was .12.

Repeated retrieval in an eyewitness paradigm has been the topic of several published reports. Shaw, Bjork, and Handal (1995) tested subjects in a typical eyewitness memory experiment in which they saw a variety of objects in a first phase. During a second phase they received practice in retrieving some objects from the display with the aid of verbal cues. In a third phase of the experiment, subjects received a free recall test on which they were to remember all of the objects they could from the original display. They found enhanced recall for those items that had received repeated retrieval (the usual testing effect), but impaired recall of the nontested objects. By implication, the act of recalling some parts of a complex event may improve retention of the recalled parts but at the cost of decreased access to other aspects of the event.

More relevant for present purposes is the research by Schooler, Foster, and Loftus (1988), who tested people repeatedly in an eyewitness memory situation on recognition tests. Subjects saw a slide sequence depicting a crime scene. For purposes of exposition, let us consider a critical detail, A, that was assessed on one or more later recognition tests. On an initial recognition test, some subjects were tested with the correct detail and a lure and were asked to choose between them (A or B?); others were tested with two lures and forced to pick between them (B or C?); and a third group simply engaged in an unrelated filler task. In the second condition above, subjects were forced to choose between two items that had not been in the original scene, and the central interest was in the effects of this forced choice on later retention. On a final recognition test, subjects were given another multiple choice test with four alternatives (Was it A, B, C, or D?).

Schooler et al. (1988) found that subjects were most accurate on the later recognition test if they had received the immediate test with the correct alternative included (A or B?), revealing another manifestation of the testing effect. However, subjects were least likely to be correct on the later test if they had been given the forced choice test for the two lures (B or C?) on the first test. Responding to the incorrect alternatives on the first test carried over to the second test. In a second experiment, Schooler et al. (1988) obtained similar results even when the test alternatives on the last test were only A and D, which indicates that it was not simply the carryover from the earlier commitment to B or C on the first test that determined performance on the final test. The outcomes of both experiments revealed that responding on a first test can impair performance on a test given later (see Hastie, Landsman, & Loftus, 1978, for similar observations).

Note that on the first recognition test, Schooler et al. (1988) directed subjects' attention to the slide sequence but presented them with incorrect alternatives. In the present experiments we also examine the effect of initial testing on later retrieval, and we also manipulate the probability with which subjects will report incorrect information on Test 1 to examine the effect of this variable on the final test. A crucial difference lies in the orientation of subjects toward the initial test. In Schooler et al's. experiments, attention was always directed toward the original event on both tests, but on Test 1 only incorrect choices were present (for some items), thus requiring an incorrect response on Test 1. In the present experiments, we manipulated test orientation on a first test. For some items, subjects were told to report only what was recalled from the original event (the slide sequence); for other items, subjects were allowed to ignore the source of

their memories and simply report an answer (i.e., they were permitted to report an answer based on what they remembered from either the slides or the narrative). On the final cued recall test, instructions were given to pay great attention to source of the memory for all questions and only to report information that had occurred in the slides. Our manipulation is similar to that of Schooler et al. (1988) in that we also encouraged subjects to report a piece of information that they would later be asked to disregard. The critical point of interest in this condition lies in whether subjects will come to believe that a wrong detail was actually present in the original event if they had reported it on a prior test. Will generating the information on a first test lead to a reality monitoring confusion on the second test, with subjects ''remembering'' an erroneous detail as actually having been presented earlier?

The present experiments were designed to address three issues. We wanted to assess (1) whether a testing effect would occur in a cued recall test in the misinformation paradigm (i.e., whether performance on a final test would differ as a function of whether or not an initial test had occurred) and (2) whether the answer to this question would differ as a function of instructions given on the first test. That is, if subjects were permitted to respond with misleading information on Test 1 in some conditions but were told to only report slide information on Test 1 in other conditions, would they differ on a final recall test? (3) The third goal was to explore the subjects' phenomenological awareness during recall of misinformation. On the second test, when they erroneously recall the item presented in the narrative, do they ''remember'' its occurrence in the slide sequence, or do they respond because the information seems familiar? The issue has also been the object of research by Lindsay (1990, 1994; Lindsay & Johnson, 1989), and Zaragoza and Koshmider (1989), among others. We asked this question by employing Tulving's (1985) remember/know procedure: Following the recall of information, subjects indicate whether they remember its occurrence in the original event or if they just know it occurred, but cannot actually remember the specific episode. A substantial body of literature attests to the usefulness of this technique (see Gardiner & Java, 1993, and Rajaram & Roediger, in press, for reviews). Zaragoza and Lane (1994) have shown that sometimes subjects report that they remember the actual occurrence of misinformation, and Roediger and McDermott (1995) also found (in a different paradigm) that subjects frequently reported remembering events that never happened (see too Payne et al., 1996).

In sum, we had three goals in conducting the two experiments reported here. First, we wanted to examine the effects of repeated testing on recall of misleading information in the eyewitness memory paradigm. Second, we attempted to manipulate production of misinformation on a first test via instructions so that we could examine the possible effects on a second test 2 days later. Third, we examined the phenomenological experience of subjects when recalling the misleading information. This was achieved by asking subjects to judge whether they remembered or knew that the information they were producing on the second test had occurred in the original event.

Both experiments had two phases. Subjects first viewed a slide sequence depicting a robbery and then read a narrative in which several items of misinformation were embedded in a generally accurate account of the slide sequence. Subjects then took a short answer test in which each item was cued as to whether the response should be made on the basis of information recalled from (a) only the slides or (b) either the slides or the narrative. We assumed (correctly) that recall of ''wrong'' (i.e., narrative) details would occur more often when subjects were permitted to recall from either source than when they were instructed to recall only from the slide sequence. Two days later subjects returned for the second test and were again asked short answer questions; this time, however, they were instructed for all items to answer questions only on the basis of what was recalled from the slide sequence. After providing each answer, subjects judged whether they *remembered* the occurrence of

the information in the slide sequence or simply *knew* that the item had occurred in the slides.

Experiment 1

Method

Subjects. The subjects were 64 Rice undergraduates who received either course credit or $5.00 for their participation. They were tested in groups from one to four.

Design. All subjects viewed a slide sequence, read a narrative that contained four items of misinformation, took a first cued recall test, and then (2 days later) took a second cued recall test. The independent variable was the instructions given on how to recall the various items on the first test; these instructions were given on an item-by-item basis. In the Either Source condition, subjects were instructed to recall the answer to the question by using information remembered from either the slide sequence or the narrative. In the Slides Only condition, subjects were told to recall the answer to the question by using only information remembered from the slide sequence (i.e., to exclude answers remembered from the narrative). In the Slides Only–Divided Attention condition, subjects were given the same instructions as in the Slides Only condition, but in addition were given a divided attention task during recall. The goal of including this condition was to reduce conscious control during recall and make responding more automatic, following the logic of Jacoby, Woloshyn, and Kelley (1989). Finally, some items were not tested in the first session. We refer to this as the No Test condition, although the items were tested 2 days later.

From the experimenters' perspective, there were four critical questions in the first phase of the experiment. The critical questions corresponded to the four items for which subjects had received misleading information. For each subject, one of the critical questions was assigned to each of the four test conditions (Either Source, Slides Only, Slides Only–Divided Attention, and No Test). The four conditions were realized within-subjects, with particular items counterbalanced across the four conditions. Two days later subjects returned and were instructed to recollect information from only the slide sequence.

In addition to these four within-subject conditions, we included a control group of 16 subjects who received the same slide sequence, the narrative (but without any misinformation), and then the first and second tests. These control subjects provided a baseline for erroneous recall when no misleading information had been provided.

Materials. The slide sequence included 33 slides taken from a longer sequence that has been used in prior research on eyewitness memory (McCloskey & Zaragoza, 1985). As described earlier, the sequence depicts a maintenance man who steals money. Four specific items were selected as critical, and narratives were prepared in the form of a police report that described many features found in the slides. Embedded in the mostly accurate account were four pieces of wrong information. For these four critical items, the detail described in the narrative differed from the actual event seen in the slides. The cued recall tests consisted of 26 short answer questions of which 4 pertained to the various items about which misinformation had been given. The 4 critical questions centered around the maintenance man's shirt sleeves (they were worn down but it was suggested that they were rolled up), the type of glasses worn by the man (actually thick plastic ones but suggested to be wire-framed), the type of soft drink on the table (Coke but suggested to be 7-Up), and the type of coffee on the shelf (Maxwell House but suggested to be Folgers). Note that the items in the slide sequence formed a different set from those used in the misinformation; they were not counterbalanced. However, our primary interest was in examining recall of suggested items on the second test as a function of the condition of the first test; items were counterbalanced over this manipulation.

Procedure. Subjects were instructed that they would see a series of slides projected onto a wall and that they should remember the details for a later test. The 33 slides por-

traying the theft were shown. Subjects were then given the police-type report and were asked to read it. They were told that it was written by a trained observer, but no mention was made as to whether the facts reported in the narrative were accurate in all their details. After reading the narrative, subjects were led down a short hall to another room and seated before one of four IBM-compatible computers. They were told that they would be given a series of questions pertaining to the robbery they had seen on the slides and that they should answer to the best of their ability in the manner specified for each question. They were then instructed in detail about the three possible test conditions: They were told that sometimes they would be instructed to recall an answer with information remembered either from the slides or the narrative (the Either Source condition); for other questions they would be asked to recall only on the basis of what they recalled from the slides (the Slides Only condition); finally, for still other questions, they would be asked to keep track of digits that appeared randomly over their headphones while they attempted to recall the answer on the basis of what they recalled from the slides (the Slides Only–Divided Attention condition). In this last condition subjects heard digits between 0 and 9 at the rate of one per second and were to track the digit series for the occurrence of three successive odd numbered digits while reading and answering the question. After answering the question, subjects were prompted to respond *yes* or *no* as to whether the sequence had contained a series of three odd-numbered digits. Finally, some items were omitted from the original test to serve as items on the No Test condition in the second test.

Altogether the test included 26 questions. Two pertained to information that was not in the slides nor in the narrative but were intended to elicit guesses. Responding on these items was quite infrequent, and they will not be discussed further. For the other 24 items, 8 were tested under each of the three instructional conditions, with conditions (Either Source, Slides Only, Slides Only–Divided Attention) assigned randomly to items. One critical test item was included in each set of 8 tested under these conditions.

Two days later subjects returned expecting to participate in another experiment. Actually, they were given the second test, which covered material presented in the previous session. Once again, subjects were given short answer cued recall questions, but this time they were carefully instructed to recall only on the basis of information remembered from the slide sequence. In addition, they were instructed to judge whether each recalled answer was *remembered* or *known*. These instructions were carefully explained, following the protocol developed in prior research (Gardiner & Java, 1990; Rajaram, 1993). Briefly, subjects were told that each time they produced an answer to a question they should indicate whether they actually remembered the occurrence of the feature in the slides (a *remember* judgment) or whether they were sure that it had occurred in the slides but did not actually remember its occurrence (a *know* judgment). Subjects answered the questions, and made

TABLE 1

PROPORTION OF DETAILS RECALLED FROM THE SLIDES OR THE NARRATIVE FOR THE FOUR CRITICAL ITEMS ON TEST 1

Recall instructions	Slides	Narrative	Intrusions	No response
Either Source	.03	.62	.05	.30
Slides Only	.18	.33	.08	.41
Slides Only–Divided Attention	.18	.34	.06	.42
Control	.34	.10	.08	.48

Note. Intrusions and nonresponses are also included.

TABLE 2

PROPORTION OF DETAILS RECALLED FROM THE SLIDES OR THE NARRATIVE ON THE FOUR CRITICAL ITEMS ON TEST 2

Condition of Test 1	Slides	Narrative	Intrusions	No response
Either Source	.13	.48	.06	.33
Slides Only	.22	.38	.03	.38
Slides Only–Divided Attention	.20	.22	.02	.56
No Test	.25	.19	.06	.50
Control (no misinformation)	.31	.09	.06	.53

Note. Intrusions and nonresponses are also included. Subjects were instructed to recall items only on the basis of memory for the slide sequence.

the remember/know judgments; they were then debriefed. The second session lasted about 20 min.

Results and Discussion

We consider the results of the two tests in turn. The data for the first and second tests appear in Tables 1 and 2, respectively. We consider performance on only the four critical items. Because recall is the dependent measure, overall responses can be decomposed into correct responding (recalling an item from the slides), recall of misleading information (recall of items from the narrative), other errors of commission (or intrusions), and failures to respond (or response omissions). We can therefore examine both correct responding and erroneous recall of information from the narrative without the two necessarily having to trade off against one another as they must do in forced choice recognition.[2] However, recall of misleading information from the narrative, indicative of false recall of the suggested information, is the primary measure of interest.

[2] At the individual item level, a trade-off necessarily occurs between information from the slides and the narrative because subjects were permitted to give at most one response. However, in the aggregate, because subjects were permitted the option to not respond, such a trade-off will not necessarily occur; that is, accurate recall and false recall do not have to vary inversely. In fact, in our results they often do not (e.g., in the results shown in Tables 2, 4, and 5).

Test 1 results. The first three rows of Table 1 show the probability of recall of information from various sources (i.e., slides, narrative, extraexperimental intrusions) and the probability of nonresponses as a function of test instructions. The final row indicates responding in the control condition, in which no misleading information had been presented. Our primary interest is in recall of misleading details from the narrative, which is shown in the second column. Not surprisingly, in the condition in which subjects were told that they could recall information from either source (the slides or the narrative), they recalled the detail from the narrative much more often than they did in the conditions in which they were told they should recall only from the slides, $F(2,126) = 10.67$, $MSE = .167$. (All levels of significance exceed the .05 level of confidence unless otherwise noted.) Two t tests confirmed that the proportion of narrative information in the Either Source condition exceeded that in the other two misleading information conditions, $t(63) = 4.08$, $SEM = .07$ and $t(63) = 3.90$, $SEM = .07$ for Slides Only and Slides Only–Divided Attention conditions, respectively.

The differences in production of the narrative information between the Either Source condition and the Slides Only conditions cannot be construed as a misinformation effect because subjects in the Either Source condition were permitted to respond with information from the narrative. However, a misinfor-

mation effect can be detected on this test by comparing recall in the Slides Only conditions with recall in the control condition. The level of false recall was considerably higher in the Slides Only conditions than in the control condition, in which subjects had received no misinformation (.33 and .34 versus .10 recalled), $t(129) = 3.79$, $SEM = .068$. In addition, a comparison of columns 1 and 2 for the Slides Only conditions shows that subjects who were instructed to recall only on the basis of the slides produced the misleading narrative items instead about twice as often as they accurately recalled information from the slides. In the control condition, the opposite pattern occurred—recall of correct information was greater than erroneous recall. Note that our control subjects used the ''misinformation response'' 10% of the time, despite the fact that they had never been exposed to this information. This occurrence indicates that the details of misinformation that we chose were plausible guesses. Still, we obtained a true misinformation effect under conditions in which subjects were instructed to exclude information from the narrative (see Lindsay, 1990 for a similar observation).

As seen in the first column of Table 1, the impact of misinformation is also apparent in the recall of information from the slides. Subjects in the control condition were almost twice as likely as those in the Slides Only conditions to provide correct information (.34 versus .18 correct recall; $t(90) = 1.73$, $SEM = .09$, and $t(88) = 1.93$, $SEM = .08$ for the Slides Only and Slides Only–Divided Attention conditions, respectively). Responding with information contained in the slides in the Either Source condition was negligible.

The condition in which subjects' attention was divided during the test showed no difference from the comparable condition with no divided attention. The intent of including this divided attention condition had been to try to block consciously controlled recall from the slides to possibly enhance false responding from the narrative. However, it was clear that this manipulation failed (although one difference appears in the results from the second test, discussed below).

Test 2 results. The primary interest in the experiment was in determining how recall of misleading information from Test 1 would carry over to Test 2. Recall that on the second test (administered 2 days later), all subjects were instructed to recall facts and details only on the basis of the slide sequence they had originally viewed, and not on the basis of memory for the narrative. Therefore, any response that included narrative information was incorrect on this test, allowing us to examine false recall of narrative information as a function of instructions given on Test 1. The data from the four critical test items are presented in Table 2 using the same four response categories as in Table 1.

Once again, primary interest centers on recall of narrative information, presented in the second column, which represents false recall. First note that in all four misinformation conditions (shown in the first four rows), subjects incorrectly recalled narrative information at higher rates than in the control condition (smallest $t(63) = 1.82$, $SEM = .07$). Therefore, misinformation led to false recall in all conditions, even when subjects were instructed to produce information only from the slide sequence. Second, the act of taking the first test increased false recall in two of the three appropriate conditions, the exception being when subjects had been engaged in a divided attention task during the first test. Responding with misinformation was greater in the Either Source condition (.48) than in the No Test condition (.19), $t(63) = 3.90$, $SEM = .07$ and also greater in the Slides Only condition (.38) than in the No Test condition (.19), $t(63) = 2.55$, $SEM = .07$. Although the misinformation effect was numerically greatest when the first test had been conducted under conditions likely to produce the wrong detail (the Either Source condition), the greater false recall in this condition than in the Slides Only condition was not statistically significant, $t(63) = 1.41$, $SEM = .07$, $p < .10$, one-tailed.

Although in many cases subjects in the Either Source condition probably knew during

the first test that they were recalling information from the narrative, when asked to recall from only the slides on the second day, they still produced the wrong details that had appeared in the narrative. This interpretation is supported by an examination of the conditional probabilities of recalling narrative material (misinformation) on the second test, given that the same item had been produced on Test 1. The conditional probability for the Slides Only condition was .76, but interestingly, the conditional probability in the Either Source condition was almost as high, at .73. The corresponding probability for the Slides Only–Divided Attention and control conditions were .64 and .20, respectively. This pattern tends to support the conclusion that source misattribution contributes strongly to the misinformation effect in eyewitness memory: information from the narrative is remembered as having occurred in the original events (Lindsay, 1990; Zaragoza & Lane, 1994; Zaragoza & Mitchell, in press; but see Zaragoza & Koshmider, 1989; Lindsay & Johnson, 1989).

We turn now to correct recall (i.e., recall of information from the slides) in Test 2; these data are shown in the far left column. As noted previously, the use of recall measures (unlike the forced choice recognition test) permits one to look for misinformation effects both in correct recall and erroneous recall without the two necessarily trading off against one another (see footnote 1). The tendency for a misinformation effect is much weaker in correct recall, with performance in the four misinformation conditions being worse, but not dramatically worse, than in the control condition. The only difference that was statistically significant was between the Either Source condition (.13 correct) and the control condition (.31), $t(63) = 3.00$, $SEM = .06$. Variation in recall within the four misinformation conditions was not significant, $F(3, 189) = 1.15$, $MSE = .157$.

A final point to note is that false recall in two of the misinformation conditions (No Test and Slides Only–Divided Attention) approximated correct recall, whereas in the Either Source and Slides Only conditions, false recall greatly exceeded accurate recall. Of course, no firm conclusions are permitted from these comparisons because different items occurred in the slides and the narrative. Still, despite instructions to recall only information from the slides, subjects readily recalled the misinformation from the narrative. We turn next to the measures of subjects' phenomenological experiences during recall.

During the second test, subjects indicated after answering each question whether they remembered the answer from the original presentation of the slide sequence or whether they simply knew that the fact requested had appeared in the slides but did not remember its actual occurrence. These data are presented in Table 3, both for correct answers (those from the slides) and for wrong answers (those from the narrative). Correct responding after 2 days was rather low (.20, on average) with remember responses (.12) occurring slightly more frequently than know responses (.08). On immediate tests, subjects usually report remembering most events that they produce, but Gardiner and Java (1991) have shown that, over time, remember responses decline and know responses increase as a proportion of total correct responses, so the data on the left side of Table 3 are not surprising. Turning to the data on the right side of Table 3, misrecall responses (.32, on average) were weighted somewhat more heavily toward remember (.19) than know (.13) responses, which is notable because subjects are claiming to remember these events as having occurred during the slide sequence when in fact the information occurred only in the narrative.

Tulving (1985) argued that tests such as cued recall that ostensibly tap episodic or explicit memory (conscious recollection) actually involve a mixture of personal (remember) and impersonal (know) experiences. Examining only remember responses permits a purer estimate of episodic recall. Accepting Tulving's (1985) assumption and examining the data on the right side of Table 3, we arrive at the conclusion that taking a prior test increases both production of misinformation responses and the probability that subjects claim to re-

TABLE 3

REMEMBER AND KNOW RESPONSES FOR CORRECT RESPONSES AND MISINFORMATION RESPONSES FOR THE FOUR CRITICAL ITEMS ON TEST 2 OF EXPERIMENT 1

	Correct (slide) responses			Misinformation (narrative) responses		
Test 1 condition	Overall	Remember	Know	Overall	Remember	Know
Either Source	.13	.09	.03	.48	.31	.17
Slides Only	.22	.09	.13	.38	.23	.14
Slides Only–Divided Attention	.20	.14	.06	.22	.16	.06
No Test	.25	.14	.11	.19	.06	.13
Control	.31	.23	.08	.09	.05	.05

member these details as having occurred in the original event. In the three conditions in which subjects had received a prior test, they reported remembering events from the slide sequence with a higher probability (.23, on average) than when no test had been given (.06), $t(190) = 3.98$, $SEM = .04$. These results show that testing can create the illusion of remembering details from an event that did not occur in the event but that were suggested later.

EXPERIMENT 2

Results of the first experiment showed (a) a strong misinformation effect in recall, both immediately after the study phase and 2 days later, (b) that the misinformation effect could be magnified by a prior test, and (c) that after 2 days subjects reported remembering the suggested detail as having occurred in the slide sequence more frequently if they had received a prior test on the detail than if they had not been tested on the detail. The purpose of Experiment 2 was to replicate these findings and to make the control condition a within-subjects rather than a between-subjects comparison, as in Experiment 1. Subjects again saw the slide sequence and read a narrative in which there were four critical items; misinformation was presented about three of them. Then they received a first test, in which instructions on how to recall items were presented with each question. The three test conditions on the first test were Either Source, Slides Only, and the Control condition in which no misinformation had been presented (and subjects were told to recall the item from the slides). Two days later they returned for a second test, in which they were asked to recall information only from the slide sequence and to make remember and know judgments about the responses. In the fourth condition, a question was asked on the second test that queried a detail for which misinformation had been presented, but it was an item that had not been previously tested (the No Test condition).

In sum, the design of Experiment 2 was like the within-subjects conditions of Experiment 1 except that the Control condition replaced the condition from Experiment 1 in which subjects had attempted to recall from slides under divided attention conditions in the first test. In addition, the entire experiment was conducted on computers, which necessitated some changes in the procedure as described below.

Method

Subjects and design. The 48 subjects were Rice University students who received course credit or summer school students who were paid $5.00 for their participation. They were tested in groups of up to 8 people. The independent variable was realized by manipulating across items whether misinformation occurred in the narrative, whether the item was tested on an initial test, and also the testing condition

for items on the first test. For one item in the slides that would later be tested, subjects received no misinformation in the narrative (the Control condition). For three other items the subjects did receive misinformation, and then they were told on Test 1 to recall information from the slides or the narrative (the Either Source condition), from only the slides (the Slides Only condition), or the item was not tested (the No Test condition). Primary interest again centered on performance on a second test given 2 days later.

Materials. The same general materials were used as in Experiment 1 (i.e., a slide sequence portraying a robbery from an office). However, each slide was transferred to a photo CD and presented on a Macintosh color monitor for 5 s. Each subject saw the same slides in the same order. Shortly after presentation of the slides, the subjects read a narrative (for 3 min) presented in the style of a police report. Among the many (mostly correct) details included in the narrative were four critical items. One item referred to one of the critical details in the slide sequence in a neutral manner, but for three other items misinformation was inserted into the narrative, thereby contradicting the nature of a detail in the original slides. The nature of the critical items was different (in two cases) from the critical items in Experiment 1, as we tried (unsuccessfully, it turned out) to replace poorer items (i.e., items not producing substantial misinformation effects) with better ones. Four versions of the narrative were created to counterbalance critical items across conditions, and 12 subjects received each version. A test containing 24 short answer questions pertaining to details in the slide sequence was developed and used on both tests.

Procedure. When subjects arrived at the lab they were seated before a Macintosh computer and told that they would see a series of pictures depicting a crime. They were told to pay particular attention to any details that might be useful in identifying the perpetrator or the crime scene as they would later be asked questions regarding them. After presentation of the pictures subjects were told that they would be given a police report to help refresh their memory for details of the event. After reading the narrative, subjects were asked to answer 18 short answer cued recall questions, which were presented in one of two random test orders—each to half the subjects. Six questions were answered under conditions in which subjects were told to recall only from the slides and to exclude responses that had occurred in the narrative (the Slides Only condition). (One question in this set pertained to an item for which subjects had received misinformation.) Twelve questions were tested under conditions in which subjects were cued, prior to reading the question, that they should recall the answer from the slides or from the narrative (the Either Source condition). (Two critical questions appeared in this set of 12, one for which subjects had received misinformation and one for which they had not.) No question was asked about the fourth misinformation item on the first test. After answering the questions, subjects were asked not to discuss the experiment with other participants and to return to the lab in 48 h for further (unspecified) experiments.

Upon their return, each subject was seated before the same computer as on Day 1, and they read instructions telling them that they would be tested again on the slide sequence they had seen 2 days ago. They were instructed only to report answers that they recalled from the slide sequence; they were explicitly told to avoid providing answers recalled only from the narrative. They were then given remember/know instructions similar to those used in Experiment 1. However, rather than the oral instructions used in Experiment 1, subjects were given written instructions on the computer in Experiment 2. Unfortunately, this mode of presentation, which we had not tried before in other experiments, was not effective. Subjects expressed confusion about the measure and the data from this judgment were not interpretable and will not be reported here.

After reading through the instructions, subjects were given 24 short answer questions pertaining to the crime witnessed in the slides,

TABLE 4

PROPORTION OF DETAILS RECALLED FROM THE SLIDES OR THE NARRATIVE ON THE FOUR CRITICAL ITEMS ON TEST 1 OF EXPERIMENT 2

Condition	Slides	Narrative	Intrusions	No response
Either Source	.33	.46	.04	.17
Slides Only	.40	.31	.04	.25
Control	.31	.21	.08	.40

Note. Also included are probability of intrusions and nonresponses.

which included six questions from the first test. The instructions reiterated that all answers given should be recalled from the slide sequence. After answering all questions, subjects were debriefed and left.

Results

We report data for the two tests separately. Unlike Experiment 1, the control condition was included as part of the within-subjects design.

Test 1 results. As shown in Table 4, subjects who were told that they could recall information from either the narrative or the slides frequently recalled the detail given in the narrative (.46 recall from the narrative, compared to .62 in Experiment 1 in this condition). However, unlike the outcome from Experiment 1, in which subjects rarely recalled the answer from the slides in this condition (.03), they rather frequently did so in Experiment 2 (.33). Thus, even though production of correct information was increased, and production of narrative information decreased in the Either Source condition relative to performance in the same condition in Experiment 1, the manipulation in Experiment 2 still succeeded in having subjects produce the misinformation quite frequently relative to the other two conditions.

An analysis of variance on recall of narrative information in the three conditions showed reliable variation, $F(2,94) = 3.85$, $MSE = .197$. In addition, recall of the narrative information was greater in the Either Source condition (.46) than in the Slides Only condition (.31), $t(47) = 1.85$, $SEM = .08$, but recall of misleading responses in the Slides Only condition did not differ reliably from that in the Control condition, $t(47) = 1.09$, $SEM = .09$. Data on correct responding (in the left column of Table 4) also show little difference among the conditions as a function of misleading information. The misleading information effect was therefore weak or absent in recall on the first test. Nonetheless, it appeared quite strongly on the delayed test 2 days later.

Test 2 results. The most direct measure of the misinformation effect is recall of the suggested detail from the narrative when subjects are instructed to exclude recall from the narrative and to recall only from the slides. The relevant data are shown in the second column of Table 5. As in Experiment 1, we found a strong misinformation effect under these conditions. Relative to the recall of such items in the control condition (.15), subjects in all three misinformation conditions showed elevated false recall, although the effect was magnified when subjects had been previously tested on the relevant items and had been induced to produce the wrong detail through instructions (.46 false recall on Test 2 after subjects had been tested in the Either Source condition on the first test). A one-way ANOVA showed significant variation among conditions, $F(3,141) = 4.04$, $MSE = .196$. Unlike the results of Experiment 1, testing in the Slides Only condition (.27) did not produce greater

TABLE 5

PROPORTION RECALLED ON TEST 2 FROM THE SLIDES OR FROM THE NARRATIVE IN EXPERIMENT 2, AS WELL AS INTRUSIONS AND NONRESPONSES

Condition of Test 1	Slides	Narrative	Intrusions	No response
Either Source	.21	.46	.00	.33
Slides Only	.52	.27	.02	.19
No Test	.15	.29	.02	.54
Control	.38	.15	.04	.44

misrecall than occurred in the No Test condition (.29). Perhaps this result is not too surprising given that no misinformation effect was found on Test 1 for this condition. Although all three misinformation conditions tended to show a misinformation effect relative to the control condition, the only significant effect occurred in the Either Source condition, $t(47) = 3.92$, $SEM = .08$.

Turning to veridical recall in the far left column of Table 5, evidence exists for a misinformation effect occurring in two of the conditions (i.e., the No Test and the Either Source conditions). Overall variation was significant, $F(3,141) = 8.12$, $MSE = .169$. Both No Test and Either Source conditions showed reliable effects of misleading information relative to the control condition, $t(47) = 2.53$, $SEM = .09$, and $t(47) = 1.94$, $SEM = .09$, respectively. Once again, the misinformation effect in the false recall measure was not well correlated with that in accurate recall, because the effect in false recall was greater in the Either Source condition than in the No Test condition, whereas in correct recall the effects were about the same.

The primary difference between Experiments 1 and 2 occurred in the Slides Only condition. Subjects in Experiment 2 were much more accurate in reporting information from only the slides, as instructed, in the first test and in ignoring misinformation during recall on the second test. We do not know why this difference occurred, but we suspect the source of the difference is performance on the first test itself. That is, subjects recalled from the slides much more accurately in Experiment 2 than in Experiment 1 on the first test. Because the act of recall promotes better recall in the future (e.g., Wheeler & Roediger, 1992), subjects were likely protected from the misinformation effect in the second test. However, this effect of testing also operated to increase the misinformation effect when the misinformation was produced on the first test in the Either Source condition, in which production of wrong information on the first test carried over to the second test. Nonetheless, the fact remains that subjects in the Slides Only condition showed a misinformation effect in Test 1 but the effect was much weaker (in false recall) or not apparent at all (in correct recall) in Experiment 2. Only further research can determine the source of this discrepancy, but it does not compromise the finding of robust misinformation effects in the Either Source condition on Test 2 in both experiments, and this effect was the main focus of this research.

General Discussion

The main findings from our experiments are that (a) a robust misinformation effect occurs in recall; (b) if subjects are induced to produce an item of misleading information on a first test, the probability is greatly increased that this information will be recalled and judged as *remembered* on a later test; and (c) subjects frequently claim to *remember* as having occurred in the visual scene details that were actually only presented later in the narrative. In addition, remembrance of the misleading information exceeded that of the information that actually occurred in the scene in some conditions. Although this outcome is confounded with specific items occurring in the slides and narrative, it is suggestive of a valid finding (especially since in other misinformation conditions, accurate recall was superior to false recall—see the Slides Only condition in Table 5). We discuss these findings and their relevance to current debates about the misinformation effect in eyewitness memory and to the role of retrieval in producing false memories.

The Misinformation Effect

Since the publication of the important analysis by McCloskey and Zaragoza (1985) of the standard paradigm developed by Loftus and her colleagues (e.g., Loftus et al., 1978) for studying the misinformation effect in eyewitness memory, researchers have debated the existence and the magnitude of the effect on various tests. The standard recognition test in which subjects are given a choice between original information and misleading information reveals large effects, whereas the modi-

fied recognition test, in which an extraexperimental distractor is paired with the original item, reveals small (but real) interfering effects (Payne, Toglia, & Anastasi, 1994). In addition, McCloskey and Zaragoza (1985) questioned whether the standard paradigm was subject to various demand characteristics and other artifacts. For example, if subjects responded with the misleading detail presented in a narrative, it might be that they were aware of that fact, were playing along with the experimenter, judged the item simply to be highly familiar, or simply believed (because the misleading information had been contained in the mostly correct narrative) that it was an accurate response.

More recently, researchers have tried to examine the issue in various ways to determine the reality of the misinformation effect when these artifactual possibilities are minimized or eliminated (see Belli, Lindsay, Gales, & McCarthy, 1994; Johnson & Seifert, 1994; Lindsay, 1990, 1994; Weingardt, Toland, & Loftus, 1994; Zaragoza & Lane, 1994; Zaragoza et al., 1996). To pick but a few examples of this work, Johnson and Seifert (1994) showed that subjects would continue to remember and to base judgments on misleading information, even after they had been told that it was wrong. (Their paradigm was different from the standard misinformation paradigm, but the work is relevant.) Weingardt et al. (1994) asked if subjects really believed their reports about misinformation by assessing their confidence in the reports with several techniques. In one approach, subjects were asked to place bets on whether the memories they reported were accurate or inaccurate. In one study subjects were willing to bet as much, on average, on false memories as on true memories. In another study, using a stricter criterion, subjects bet less on memories from suggested information than on the events that actually happened, but nonetheless some subjects bet the maximum amount possible that the memories of suggested information had actually occurred.

In a study that is directly related to the present work, Belli et al. (1994) tested subjects in the standard misinformation paradigm, except at the time of test subjects were given a cued recall test and warned that some information in the narrative had inaccurately represented the actual events. They were also given two questions per item; one question asked subjects to report the answer that had been contained in the slides, and the other question asked subjects to report what had been contained in the narrative. Subjects were less likely to recall correct answers if they had been exposed to misleading information; they also frequently recalled the details suggested in the narrative as having occurred in the slide sequence. Similar results were reported by Zaragoza and Lane (1994) using a recognition test; in addition, both these findings appear in the experiments reported here.

The other set of experiments directly related to our own use the logic of opposition introduced by Jacoby et al. (1989), in which subjects are told to exclude certain information from their recollections. In the context of eyewitness experiments, Lindsay (1990) had subjects participate in a standard misinformation experiment, but at the time of the test, he told subjects that the narrative did not contain any correct answers to the questions given and that they should recall only from the slides. (They were told that if a detail was remembered from the narrative, it was wrong.) When subjects saw the slides, received the narrative immediately thereafter, and then were tested 2 days later, they often reported the suggested details as having actually occurred in the event. Weingardt et al. (1994) reported similar observations with a different test and slightly different test instructions.

The present experiments embed elements of these other experiments—cued recall tests and instructions to recall only from the slides and not from the narrative—while adding the use of remember/know judgments to determine if subjects remember the suggested events as occurring in the slide sequence (see too Zaragoza & Mitchell, in press). On the second (delayed) test, in which we were most interested, subjects were told to recall only from the slide sequence (i.e., not from the

narrative) and to judge, for each answer, whether they remembered the occurrence in the slides or simply knew that the detail had been in the slide sequence. In both experiments, we found a large misinformation effect on the delayed test in the condition in which subjects had produced information from either source on the first test. In addition, if we consider *remember* responses a purified form of episodic recall, with responding based on familiarity or general knowledge eliminated, then subjects remembered misinformation more often if they had been previously tested than if they had not.

Although misinformation effects did not occur in all measures in each of our experiments, the effect was quite robust overall, especially when subjects were tested twice. False recall levels were high (sometimes exceeding those of accurate recall), and the effect was magnified by repeated testing, as discussed below. The results based on correct recall were somewhat more variable, but nonetheless tended to show misinformation effects (except in the first test of Experiment 2). Using recall as the measure permits a more fine-grained analysis of performance, because subjects are not forced to respond to every test cue, as Lindsay (1994) has noted. In addition, using various instructions to recall or recognize information either only from the slides, from the narrative, or from either or both sources also permits more information to be gained (see Zaragoza & Lane, 1994, among others). The instruction in the second (critical) test of our experiments directed subjects to recall information only from the slide sequence. Although this instruction does not fully instantiate the logic of opposition as used in Lindsay (1990), in which subjects were told that if they remembered a detail from the narrative, they could be sure that it was wrong, our instruction seems more plausible in trying to generalize results beyond the laboratory. Rarely can one be assured, outside experiments embodying the logic of opposition, that any event-related information received after the original occurrence is incorrect. Rather, people are usually asked (say, in courtroom testimony) to remember accurately what they can from a specified event. Use of our instruction to try to recall only from the slides, combined with the use of *remember* judgments, would appear to overcome demand characteristics and more closely approximate testimony in legal settings.

The Role of Repeated Production in Creating Memory Illusions

The primary manipulation of interest in our experiments was the condition in which subjects were encouraged during a first test to report information either from the narrative or from the slides and then required to recall only information from the slides on a second test. On the first test they frequently produced information from the narrative, especially in Experiment 1, probably because this information had been provided more recently than that in the slides. Many of these items produced on the first day were also recalled 2 days later even though subjects were instructed to recall only information from the slides. In addition, subjects frequently claimed to *remember* the moment of occurrence of these items that had actually only been suggested in the narrative. The use of remember/know judgments permit assessment of true ''false remembering'' in a way that most other types of judgments do not (see Lindsay (1994) for a critique of other techniques). Roediger and McDermott (1995), Payne et al. (1996), and Zaragoza and Mitchell (in press) have also obtained false remembering using this technique.

Misinformation is usually presented from external sources. Our technique of having subjects produce the misinformation themselves after it had been presented externally led to more powerful misinformation effects than in other conditions. This outcome fits well with the source monitoring approach to the misinformation effect (Lindsay & Johnson, 1989; Lindsay, 1994). Previous experiments have shown that imagining events causes confusion as to whether the event was only read about or was actually carried out (Finke, Johnson, & Shyi, 1988; Johnson, Raye, Foley & Foley, 1981; Lindsay, 1990). Similarly, in our experiments, actually producing the information in

the context of a test leads to the attribution, on a later test, of *remembering* the information.

Zaragoza and Mitchell (in press) and Mitchell and Zaragoza (1996) found that repeating misinformation enhances its impact, increasing misinformation effects. Their results involved external repetitions, whereas ours involved having subjects themselves reproduce the misinformation on a first test. However, both techniques led to similar results in that the misinformation effect is greater after repeated exposure to the misinformation. A topic for future research is whether internal generation creates greater interference and false recall than does external presentation.

Further evidence that response production on a first test can create confusion on a later test comes from experiments in which subjects are encouraged to guess during recall. Hastie et al. (1978) encouraged subjects to guess on a first test in an eyewitness memory paradigm. Relative to control subjects, those who had guessed on a first test were more likely to make errors on a later test. In addition, subjects were often more confident about their guesses on a second test than on a first test. Roediger, Wheeler, and Rajaram (1993) and Jacoby, Kelley, and Dywan (1989) report additional evidence for how guessing can undermine accurate recollection. Other experiments have also shown the power of false recall on a first test to enhance erroneous recall later (e.g., McDermott, 1996). Of course, Schooler et al. (1988) reported a similar effect on successive recognition tests, as reviewed in the Introduction.

The general point is that when a person is queried about memory for a distant event, recall is almost certain to be influenced not only by what transpired during the event in question, but by all the previous recollections of that event. The information retrieved from the most recent account of the event may be a more powerful determinant of the current recollection than the original event itself.

REFERENCES

ANDERSON, M., & NEELY, J. H. (in press). In E. L. Bjork and R. A. Bjork (Eds.), *Handbook of Perception and Cognition,* Vol. 10. New York: Academic Press.

BELLI, R. F., LINDSAY, D. S., GALES, M. S., & MCCARTHY, T. T. (1994). Memory impairment and source misattribution in postevent misinformation experiments with short term retention intervals. *Memory & Cognition,* **21,** 40–54.

DARLEY, C. F., & MURDOCK, B. B. (1971). Effects of prior free recall testing on final recall and recognition. *Journal of Experimental Psychology,* **91,** 66–73.

DENT, H., & STEPHENSON, G. M. (1979). An experimental study of the effectiveness of different techniques of questioning child witnesses. *British Journal of Social and Clinical Psychology,* **18,** 41–51.

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

FINKE, R. A., JOHNSON, M. K., & SHYI, G. C.-W. (1988). Memory confusions for real and imagined completions of symmetrical visual patterns. *Memory & Cognition,* **16,** 133–137.

GARDINER, J. M., & JAVA, R. I. (1991). Forgetting in recognition memory with and without recollective experience. *Memory & Cognition,* **19,** 617–623.

GARDINER, J. M., & JAVA, R. I. (1993). Recognizing and remembering. In A. Collins, M. A. Conway, S. E. Gathercole, & P. E. Morris (Eds.), *Theories of memory* (pp. 163–188). Hillsdale, NJ: Erlbaum.

GARDINER, J. M., & JAVA, R. I. (1990). Recollective experience in word and nonword recognition. *Memory & Cognition,* **18,** 23–30.

HASTIE, R., LANDSMAN, R., & LOFTUS, E. F. (1978). Eyewitness testimony: The dangers of guessing. *Jurimetrics Journal,* **19,** 1–8.

HOGAN, R. M., & KINTSCH, W. (1971). Differential effects of study and test trials on long-term retention and recall. *Journal of Verbal Learning and Verbal Behavior,* **10,** 562–567.

JACOBY, L. L., KELLEY, C. M., & DYWAN, J. (1989). Memory attributions. In H. L. Roediger & F. I. M. Craik (Eds.), *Varieties of memory and consciousness: Essays in honour of Endel Tulving* (pp. 391–422). Hillsdale, NJ: Erlbaum.

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

JOHNSON, H. M., & SEIFERT, C. M. (1994). Sources of the continued influence effect: When misinformation in memory affects later inferences. *Journal of Experimental Psychology: Learning, Memory, and Cognition,* **20,** 1420–1436.

JOHNSON, M. K., HASHTROUDI, S., & LINDSAY, D. S. (1993). Source monitoring. *Psychological Bulletin,* **114,** 3–28.

JOHNSON, M. K., RAYE, C. L., FOLEY, H. J., & FOLEY, M. A. (1981). Cognitive operations and decision bias in reality monitoring. *American Journal of Psychology,* **94,** 37–64.

JONES, T. C., & ROEDIGER, H. L. (1995). The experiential basis of serial position effects. *European Journal of Cognitive Psychology,* **7,** 65–80.

LINDSAY, D. S. (1990). Misleading suggestions can impair eyewitnesses' ability to remember event details. *Journal of Experimental Psychology: Learning, Memory, and Cognition,* **16,** 1077–1083.

LINDSAY, D. S. (1994). Memory source monitoring and eyewitness testimony. In D. F. Ross, J. D. Read, and M. P. Toglia. *Adult eyewitness testimony: Current trends and developments* (pp. 27–55). New York: Cambridge Univ. Press.

LINDSAY, D. S., & JOHNSON, M. K. (1989). The reversed eyewitness suggestibility effect. *Bulletin of the Psychonomic Society,* **27,** 111–113.

LOCKHART, R. S. (1975). The facilitation of recognition by recall. *Journal of Verbal Learning and Verbal Behavior,* **14,** 253–258.

LOFTUS, E. F. (1979a). *Eyewitness testimony*. Cambridge, MA: Harvard Univ. Press.

LOFTUS, E. F. (1979b). The malleability of memory. *American Scientist,* **67,** 312–320.

LOFTUS, E. F. (1993). Made in memory: Distortions in recollection after misleading information. In D. L. Medin (Ed.). *The psychology of learning and motivation: Advances in theory and research* (pp. 187–215). New York: Academic Press.

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.

MCCLOSKEY, M., & ZARAGOZA, M. S. (1985). Misleading postevent information and memory for events: Arguments and evidence against memory impairment hypotheses. *Journal of Experimental Psychology: General,* **114,** 1–16.

MCDANIEL, M. A., KOWITZ, M. D., & DUNAY, P. K. (1989). Altering memory through recall: The effect of cue-guided retrieval processing. *Memory & Cognition,* **17,** 423–434.

MCDERMOTT, K. B. (1996). The persistence of false memories in list recall. *Journal of Memory and Language,* **35,** 212–230.

MITCHELL, K. J., & ZARAGOZA, M. S. (1996). Repeated exposure to suggestion and false memory: The role of contextual variability. *Journal of Memory and Language,* **35,** 246–260.

PAYNE, D. G. (1987). Hypermnesia and reminiscence in recall: A historical and empirical review. *Psychological Bulletin,* **101,** 5–27.

PAYNE, D. G., ELIE, C. J., BLACKWELL, J. M., & NEUSCHATZ, J. S. (1996). Memory illusions: Recalling, recognizing, and recollecting events that never occurred. *Journal of Memory and Language,* **35,** 261–285.

PAYNE, D. G., TOGLIA, M. P. & ANASTASI, J. S. (1994). Recognition performance level and the magnitude of the misinformation effect in eyewitness memory. *Psychonomic Bulletin & Review,* **1,** 376–382.

POOLE, D. A., & WHITE, L. A. (1991). Effects of question repetition on the eyewitness testimony of children and adults. *Developmental Psychology,* **27,** 975–986.

POOLE, D. A., & WHITE, L. A. (1993). Two years later: Effects of question repetition and retention interval on the eyewitness testimony of children and adults. *Developmental Psychology,* **29,** 844–853.

POOLE, D. A., & WHITE, L. T. (1995). Tell me again and again: Stability and change in the repeated testimonies of children and adults. In M. S. Zaragoza, J. R. Graham, G. C. N. Hall, R. Hirschman, & Y. S. Ben-Porath (Eds.), *Memory and testimony in the child witness* (pp. 24–43). Thousand Oaks, CA: Sage.

RAJARAM, S. (1993). Remembering and knowing: Two means of access to the personal past. *Memory & Cognition,* **21,** 89–102.

RAJARAM, S., & ROEDIGER, H. L. (in press). Remembering and knowing as states of consciousness during retrieval. J. D. Cohen & J. W. Schooler (Eds.) *Scientific approaches to consciousness.* Hillsdale, NJ: Erlbaum.

ROEDIGER, H. L. (1974). Inhibiting effects of recall. *Memory & Cognition,* **2,** 261–269.

ROEDIGER, H. L., & GUYNN, M. J. (in press). Retrieval processes. In E. L. Bjork and R. A. Bjork (Eds.), *Handbook of perception and cognition,* Vol. 10. New York: Academic Press.

ROEDIGER, H. L., & MCDERMOTT, K. B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition,* **21,** 803–814.

ROEDIGER, H. L., & NEELY, J. H. (1982). Retrieval blocks in episodic and semantic memory. *Canadian Journal of Psychology,* **36,** 213–242.

ROEDIGER, H. L., & SCHMIDT, S. R. (1980). Output interference in the recall of categorized and paired associate lists. *Journal of Experimental Psychology: Human Learning and Memory,* **6,** 91–105.

ROEDIGER, H. L., WHEELER, M. A., & RAJARAM, S. (1993). Remembering, knowing, and reconstructing the past. In D. L. Medin (Ed.) *The Psychology of Learning and Motivation: Advances in research and theory,* (pp. 97–134). Academic Press: San Diego.

SCHACTER, D. L., VERFAELLIE, M., & PRADERE, D. (1996). The neuropsychology of memory illusions: False recall and recognition in amnesic patients. *Journal of Memory and Language,* **35,** 319–334.

SCHOOLER, J. W., FOSTER, R. A., & LOFTUS, E. F. (1988). Some deleterious consequences of the act of recollection. *Memory & Cognition,* **16,** 243–251.

SCRIVENER, E., & SAFER, M. A. (1988). Eyewitnesses show hypermnesia for details about a violent event. *Journal of Applied Psychology,* **73,** 371–377.

SHAW, J. S., & BJORK, R. A., & HANDEL, A. (1995). Retrieval-induced forgetting in an eyewitness-mem-

ory paradigm. *Psychonomic Bulletin & Review,* **2,** 249–253.

SMITH, A. D. (1971). Output interference and organized recall from long-term memory. *Journal of Verbal Learning and Verbal Behavior,* **10,** 400–408.

SPITZER, H. F. (1939). Studies in retention. *Journal of Educational Psychology,* **30,** 641–656.

THOMPSON, C. P., WENGER, S. K., & BARTLING, C. A. (1978). How recall facilitates subsequent recall: A reappraisal. *Journal of Experimental Psychology: Human Learning and Memory,* **4,** 210–221.

TULVING, E. (1985). Memory and consciousness. *Canadian Psychologist,* **26,** 1–12.

TULVING, E. & ARBUCKLE, T. Y. (1963). Sources of intratrial interference in paired-associate learning. *Journal of Verbal Learning and Verbal Behavior,* **1,** 321–334.

WEINGARDT, K. R., TOLAND, H. K., & LOFTUS, E. F. (1994). Reports of suggested memories: Do people truly believe them? In D. F. Ross, J. D. Read, and M. P. Toglia. *Adult eyewitness testimony: Current trends and developments* (pp. 3–26). New York: Cambridge Univ. Press.

WHEELER, M. A., & ROEDIGER, H. L. (1992). Disparate effects of repeated testing: Reconciling Ballard's (1913) and Bartlett's (1932) results. *Psychological Science,* **3,** 240–245.

ZARAGOZA, M. S., & KOSHMIDER, J. W. (1989). Misled subjects may know more than their performance implies. *Journal of Experimental Psychology: Learning, Memory, and Cognition,* **15,** 246–255.

ZARAGOZA, M. S., & LANE, S. (1994). Source misattributions and the suggestibility of eyewitness memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition,* **20,** 934–945.

ZARAGOZA, M. S., & MITCHELL, K. J. (in press). Repeated suggestion and the creation of false memories. *Psychological Science.*

(Received July 24, 1995)
(Revision received December 23, 1995)