

Creating new memories that are quickly accessed and confidently held

ELIZABETH F. LOFTUS, KAREN DONDERS, and HUNTER G. HOFFMAN
University of Washington, Seattle, Washington

and

JONATHAN W. SCHOOLER
University of Pittsburgh, Pittsburgh, Pennsylvania

In two experiments involving a total of 542 subjects, a series of slides depicting a burglary was shown. After the initial event, subjects were exposed to one or more narratives about the event that contained some misinformation or neutral information about four critical details. Finally, subjects were tested on their memories of what they saw, and their reaction times and confidence levels were measured. When subjects took a standard test in which the misinformation item was a possible response option, they responded very quickly and confidently when making this incorrect choice. Misled subjects responded as quickly and confidently to these "unreal" memories as they did to their genuine memories. It does not seem, then, that the misinformation effect arises from a large proportion of subjects who must resolve a conflict between two memories when they are tested, a conflict that would be expected to take time. When subjects took a modified test in which the misinformation item was not a possible response, misled subjects were as accurate as were controls, but they responded more slowly, regardless of whether they ultimately chose the right or wrong option. These findings indicate that misinformation does introduce some form of interference not detected by a simple test of accuracy.

Information presented after an event can change a person's report of that event. When the new information is misleading, it can produce errors in that report. Stop signs are recalled as yield signs (Loftus, Miller, & Burns, 1978), hammers are recalled as screwdrivers (Belli, 1989; McCloskey & Zaragoza, 1985), traffic signs are recalled as stop signs (Wagenaar & Boer, 1987), Coke cans are recalled as cans of Planter's peanuts (Zaragoza, McCloskey, & Jamis, 1987), *Vogue* magazine is remembered as *Mademoiselle* (Tversky & Tuchin, 1989), green plastic pictures are remembered as yellow (Belli, 1988), breakfast cereal is remembered as eggs (Ceci, Ross, & Tolia, 1987a, 1987b; Ceci, Tolia, & Ross, 1988), a clean-shaven man is recalled as having a mustache (Gibling & Davies, 1988), the word *Yukon* on a T-shirt is remembered as *Nixon* (Sheehan & Tilden, 1986), and a man named Dr. Henderson is remembered as being named Dr. Davidson (Geiselman, 1988). We refer to the misreporting of information after exposure to misleading information as the *misinformation* effect. In numerous laboratories, the misinformation effect has been obtained, and there seems to be little doubt that erroneous reporting is easy to induce.

The misinformation effect has stimulated a significant degree of interest from memory researchers, not only in the United States, but also in Canada, Great Britain, Germany, Australia, and The Netherlands. Most of the research (e.g., Brainerd & Reyna, 1988; Gudjonsson, 1986; Hammersley & Read, 1986; Kohnken & Brockmann, 1987; Kroll & Ogawa, 1988; Morton, Hammersley, & Bekerian, 1985; Register & Kihlstrom, 1988; Smith & Ellsworth, 1987) has used a three-stage procedure in which subjects first experience an event, then receive new information about the event, and, finally, take a test of memory for the event. Suppose a subject saw a hammer as part of a critical event, and then received misinformation about a screwdriver. In the traditional testing procedure, the subject could respond with the misinformation. If the subject saw a hammer, later heard about a screwdriver, the traditional (or *standard*) test might ask whether the subject saw a hammer or a screwdriver. The typical result is that subjects are significantly less accurate after misinformation than they are in the absence of misinformation. Put another way, only about 40% of misled subjects might correctly choose "hammer," but perhaps as many as about 70% of control subjects might correctly choose "hammer."

A question arises as to why subjects choose "screwdriver" on the standard test. There are a number of possible reasons. Some subjects might have failed to encode the hammer in the first place, and they choose "screwdriver" on the test because they remember reading about it in the postevent material. Some subjects might remem-

This research was supported by grants from the National Institute of Mental Health and the National Science Foundation. We thank William Kohlmeyer and Monica Van Hoomissen for their assistance in running the experiments. Robert Belli provided enormously helpful comments on an earlier draft of the paper. Please address all correspondence to Elizabeth Loftus, Department of Psychology, University of Washington, Seattle, Washington 98195.

ber both the hammer and the screwdriver, and they choose "screwdriver" on the test because they trust the experimenter's materials more than they trust their own memories. Some subjects might remember the postevent information and might simply be confused about its source, misattributing information from one source to another (Johnson, 1987; Lindsay & Johnson, 1987). Finally, it is possible that, for some subjects, the "screwdriver" information could have changed the earlier memory of hammer. In a review of these reasons, Wells and Turtle (1987) state that it "remains unclear whether subjects' memories are actually altered by misleading [information] or whether their reports are altered while their memories remain intact and unaltered" (p. 373). Later, they argue that "although there seems little doubt that misleading postevent information distorts the testimony of many respondents, it is not clear if it actually alters the respondents' original memories" (p. 374). The determination of circumstances under which these various processes are operating is critical not only for theoretical models of forgetting, but also for applied questions in the domains of eyewitness testimony and decision making (Johnson & Hasher, 1987).

Still another explanation exists. Recall that in the typical misinformation study, a person sees one object and later learns, via postevent presentation, that it was a different object. In at least some individuals, such a state of affairs may produce a conflict that needs to be resolved. How is this conflict resolved? Does conflict resolution occur at the moment the new information (about, for example, a screwdriver) is introduced? Or does the resolution take place only when the person is finally asked to recall the tool that was lifted from the tool box (hammer vs. screwdriver)? Put another way, is the conflict resolved at the time the postevent information is comprehended, or is it resolved at the time of the test? Measurements of the speed with which subjects answer test questions bear on this issue. If a person resolves a conflict while answering the test question, response times (RTs) would be expected to be longer following exposure to misinformation than when misinformation was never encountered. On the other hand, if any conflict is resolved prior to the test, RTs following misinformation should not be longer than when following neutral information.

Past research examining the speed of responding to test questions (e.g., see Cole & Loftus, 1979; Tousignant, Hall, & Loftus, 1986) provided preliminary evidence that misinformation did not lead subjects to respond more slowly to test questions. If anything they were slightly faster. Such findings have been used to argue against the notion that many subjects enter into the test with two memory representations, one for the original item (e.g., hammer) and one for the misleading item (e.g., screwdriver), and the conflict between them is resolved during the testing period. However, the results are more consistent with the notion that subjects are not resolving a conflict during the test period, or, alternatively, if a conflict

between two memories did exist at all, it was resolved prior to the measured RT.

Past research examining the speed of responding to test items after misinformation failed to conditionalize on whether the subject was ultimately correct or incorrect, so potentially valuable information was not obtained. What if misinformation was associated with very slow correct responses (choice of the original item, "hammer") and very fast incorrect responses (choice of the misinformation option, "screwdriver"). Averaging would lead to an incomplete picture of how misinformation influences the speed of responding.

The hypothesis that misinformation might differentially affect the speed of making correct versus incorrect responses is a reasonable one. Consider the speed with which control subjects respond to the test between "hammer" and "screwdriver." Those who correctly choose "hammer" might be expected to do so more quickly than would those who incorrectly chose "screwdriver." We expect this result because the incorrect choices reflect pure guessing, whereas the correct choices reflect a combination of guessing, on the part of some subjects, and genuine memory, on the part of others.

What would misinformation do to the relationship between accuracy and RT? Would misled subjects who are correct also respond more quickly than would those who err? In fact, if misled subjects who err do so because they have what feels to them like a genuine memory for the misinformation, they may respond quite quickly, perhaps at least as fast, or even faster when wrong than when right. On the other hand, if misled subjects who err do so because during the testing period they resolve a conflict between two memories in favor of the misinformation option, they may respond quite slowly.

Experiment 1 examined RT at the time of final testing. A series of slides depicting a residential burglary was shown. After the initial event, the subjects were exposed to a narrative about the event that contained some misinformation or neutral information about certain critical details. Finally, the subjects were tested on their memories of what they had seen, and their RTs were measured. The test was between the original item and the postevent item (the standard test). The comparison of primary interest was between RT following misleading information and RT following neutral information. We were particularly interested in whether misled subjects who err (i.e., choose the misinformation option) tend to do so quickly or not.

Experiment 1 included another testing condition: One group of subjects was not permitted to choose the misinformation item on the test. We used a testing procedure adopted by McCloskey and Zaragoza (1985) in which the original item (e.g., "hammer") is paired for testing purposes with a new item (e.g., "wrench") that was never seen nor presented as misinformation. The misinformation item (e.g., "screwdriver") is not an option on the test. How would subjects given this *modified* test behave in terms of confidence and speed of responding? Control

subjects should respond as they ordinarily would—faster when correct than when incorrect. But when given misinformation and not allowed to choose it, the subjects who would have been fast and confident about the erroneous misinformation response are now going to face a conflict. They go into the testing situation with a propensity to choose the misinformation item, but they can't do this. The modified test ("hammer" vs. "wrench") essentially tells them that the preferred choice is wrong and that they must choose something else. Resolving this conflict is bound to take time.

Setting aside the issue of alteration of underlying memory traces, we can still ask about the nature of the individual's report about his or her memory. We know a great deal about the circumstances under which people can be induced to report erroneously, yet we know little about the detailed characteristics of those reports. Several past studies have shown that the confidence of reports based upon misinformation can be quite high (e.g., Greene, Flynn, & Loftus, 1982; Tversky & Tuchin, 1989), and the erroneously reported items can be described in great detail (Schooler, Clark, & Loftus, 1988; Schooler, Gerhard, & Loftus, 1986). Yet, a real understanding of the nature of these "suggested" memories would benefit from further examination of their specific characteristics.

EXPERIMENT 1

Method

The procedure used in Experiment 1 was similar to that typically used in misinformation experiments, except for the fact RTs were measured while subjects were tested.

Subjects. The subjects were 204 students from the University of Washington who received class credit for their participation. They were tested in groups ranging in size from 2 to 6 persons.

Stimuli. The slide sequence and postevent narrative were adapted from McCloskey and Zaragoza's (1985) study. The 79 slides depicted an incident in which a man goes into an ordinary-looking office, ostensibly to repair a broken chair. While inside, he snoops around, steals \$20 and a calculator, and leaves. The sequence contains four critical slides, and, for each critical slide, there were three different versions. The critical items were a coffee jar, a magazine, a soft drink, and a tool. The three critical versions of each item were: "Folgers," "Nescafé," "Maxwell House"; "Glamour," "Vogue," "Mademoiselle"; "Coke," "7UP," "Sunkist"; and "hammer," "screwdriver," "wrench." Approximately one third of the subjects saw each version of a critical item.

The narrative presented to the subjects after the event contained 735 words. It accurately described the details of the event, with the exception of the critical items. For a given subject, the narrative mentioned two critical items in a misleading way and two in a neutral way. For example, for half of the subjects who actually saw a hammer, their narrative referred to a "tool"; for the other half, their narrative referred to either a "screwdriver" or a "wrench." For each version of each critical item, the alternative versions were used about as often as was the misleading information.

Procedure. The subjects were informed that the experiment was concerned with their intuitions about memory. They were told that they would see a slide sequence depicting an event and would then read a written description of the event. As a pretense for their task, they were told that they would have to judge whether memory for an event would generally be better with the visual or the verbal mode of presentation. The slides were presented at a rate of

5 sec/slide. After a 10-min filler activity, the subjects read a narrative about the slides. To give the narrative some credibility, the subjects were told that the narrative had been written by a professor who had watched the slides very carefully.

After another 10-min filler activity, the subjects took their final test. The test consisted of a 36-item forced-choice recognition test similar to that used by McCloskey and Zaragoza (1985). Every item on the test was of the same form—a sentence with a missing word and two alternatives. Of the 36 test items, 32 were fillers and 4 were critical items. The two alternatives presented depended upon whether the subject was being given a standard test or a modified test. Fifty-one subjects took the standard test, and 153 subjects took the modified test. (As McCloskey and Zaragoza did, we used fewer subjects in the standard test group than in the modified test group since the misleading effect is generally so large in the standard testing condition.)

Each subject took the test while seated at an IBM computer. For each test item, the subject first read the sentence with the missing word (e.g., "The man slid the calculator beneath a _____ in his tool box"). After reading the question, he/she pressed the space bar that initiated the test phase 1.5 sec later. The 1.5-sec interval was included in order to give the subject time to move his/her fingers from the space bar to the response keys. The left hand moved to the 1 key and the right hand moved to the 0 key. In the test phase, each subject was presented with two alternatives and had to decide, as fast as possible, which alternative correctly corresponded to what he/she had seen in the slide sequence. The subject responded by pressing the 1 key or the 0 key. After every question, each subject entered his/her confidence about that question on a 5-point rating scale, where 1 = *not at all confident* and 5 = *very confident*. The next test item followed 3.5 sec after the confidence response. In the set of 36 items, Questions 8, 15, 23, and 31 contained the critical items.

Results

Standard test. Not surprisingly, the subjects who took the standard test were significantly more accurate on control items than they were on misled items. Mean performance was 68% correct for control items and 32% correct for misled items—a difference of 36% [$t(50) = 5.01$, $p < .001$]. Having established that a significant misinformation effect was obtained with these materials, we next examined the RT data.

Figure 1 (left panel) shows the RT data on the standard test. These data show no overall tendency for the subjects to respond more slowly in the face of misinformation. Thus, we have no evidence to support the hypothesis that misinformation produces a conflict that must be resolved during the measured RT. Moreover, on control items, the subjects were faster when they were correct than when they were incorrect; however, on misled items, this relationship was reversed.

Statistical analyses on the RT data were carried out on 14 macrosubjects created by combining data from subjects who were tested at the same time. A macrosubject analysis was used because it increased stability of measurement and eliminated empty cells in the analysis. We briefly explain these considerations.

Stability of measurement was a problem due to the small number of observations per condition. There were four critical slides shown to each subject and four corresponding test trials on which response measures were collected. On a given test trial, furthermore, a subject could respond either correctly or incorrectly. It should be apparent that

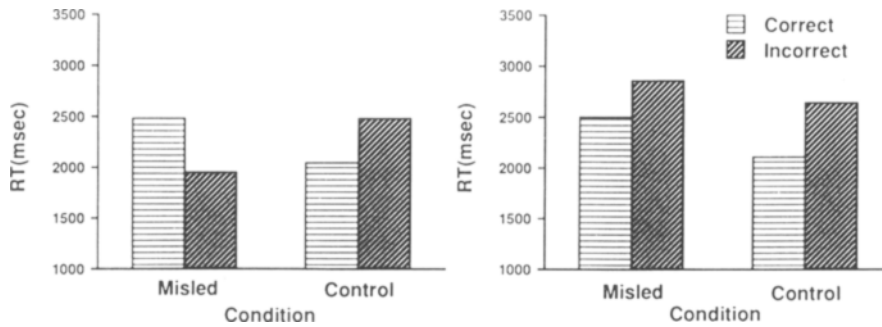


Figure 1. Mean reaction times obtained in Experiment 1. Left panel: standard test. Right panel: modified test.

the number of observations per condition for a single subject was small. By combining data across groups of several subjects, we eliminated empty cells in the final data matrix, and RT averages could be based on a larger number of observations.

The significance of the differences between these conditions was tested using a 2×2 within-subject analysis of variance. Overall, the subjects were equally fast on misled and control items [$F(1,13) < 1$] and equally fast with their correct and incorrect answers [$F(1,13) < 1$]. However, there was a highly significant interaction [$F(1,13) = 9.77, p < .01$].

Figure 2 (left panel) shows the confidence data on the standard test. These data indicate that the subjects responded to misled items somewhat more, not less, confidently than they did to control items. We can also see in the data that on control items, the subjects were more confident about their correct answers than they were about their incorrect ones. However, on misled items, this relationship between confidence and accuracy disappeared.

Statistical analyses on the confidence data were carried out on the 14 macrosubjects created by combining data for groups of approximately 4 subjects. A 2×2 within-subject analysis of variance revealed that the subjects were not significantly more confident on misled items than they were on control items [$F(1,13) = 2.6, .10 > p > .15$]. They were, however, significantly more confident when they were correct than when they were incorrect [$F(1,13) =$

$7.65, p < .02$]. The interaction was marginally significant [$F(1,13) = 4.03, p = .06$].

In sum, when the subjects took the standard test, in which they could select the misinformation item if they wished to, they tended to select it quickly and with a high degree of confidence. We now turn to the question of what happens when subjects are not permitted to select the misinformation item.

Modified test. The subjects who took the modified test were not significantly more accurate on control items than they were on misled items. Mean performance was 64% correct for control items and 68% correct for misled items [$t(152) = 1.05, p > .25$]. The lack of misinformation effect with the modified test is consistent with what McCloskey and Zaragoza (1985) found using these materials.

Figure 1 (right panel) shows that, on the modified test, the subjects responded faster to control items than they did to misled items. In addition, the subjects were faster when they were correct than when they were incorrect, both for control items and for misled items.

Statistical analyses on the RT data were carried out on 36 macrosubjects created by combining data from subjects who were tested at the same time. A 2×2 within-subject analysis of variance on the macrosubject data revealed that the subjects were faster on control items than they were on misled items [$F(1,35) = 9.34, p < .01$] and faster when they were correct than when they were in-

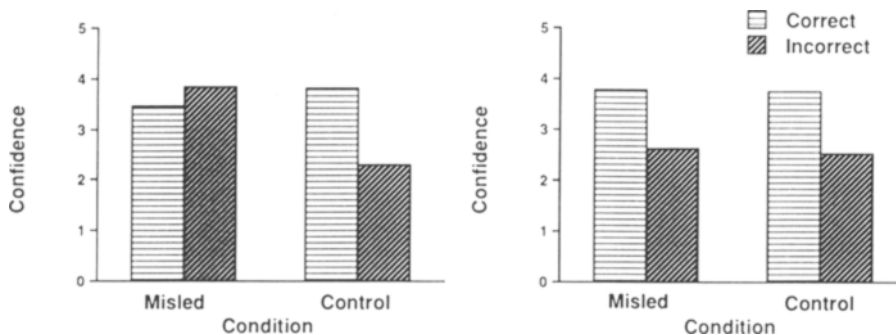


Figure 2. Mean confidence ratings obtained in Experiment 1. Left panel: standard test. Right panel: modified test.

correct [$F(1,36) = 8.65, p < .01$]. The interaction was not significant ($F < 1$).

Figure 2 (right panel) shows the confidence data on the modified test. The subjects were more confident when correct than when incorrect, both for misled and control items. An analysis of variance on the 36 macrosubjects revealed that the subjects were more confident when correct than when incorrect [$F(1,36) = 103.67, p < .0001$], whereas no other main effects or interactions were significant ($F < 1$).

Discussion

The chief results from Experiment 1 can be summarized quite readily. When the subjects were given the modified test, they, of course, could not choose the suggested item. They were forced to choose between two other options—the one they originally saw and a completely novel item. The misled and the control subjects were equally accurate in terms of choosing the original item, but it took the misled subjects significantly longer to do so. We discuss the implications of this result after considering a second experiment.

When the subjects were given the standard test, permitting them to choose the suggested item, they chose that item quickly and confidently. This finding argues against the hypothesis that subjects who have original information and misinformation must resolve a conflict when they are tested, a conflict that would be expected to take time.

One nagging problem with this interpretation remains: Subjects may resolve conflicts, if any exist, *before* the alternatives appear. These processes may occur during the presentation of the sentence frames and/or during the 1.5-sec delay between pressing the space bar and the actual appearance of the alternatives. Therefore, a second experiment was conducted. In Experiment 2, the 1.5-sec delay was eliminated so that the subjects could not use this period to anticipate choices or to engage in other conflict-resolution activities. Moreover, new instructions placed pressure on the subjects to reduce the likelihood of any extraneous mental processes occurring during the presentation of the sentence frames.

EXPERIMENT 2

One disturbing aspect of Experiment 1 is that the performance of the control group was not particularly good, with about 65% correct where chance was 50%. When the performance of a control group is relatively poor, it means that the opportunity for observing a misinformation effect (e.g., impaired performance for misled subjects) is limited. Thus, we replicated Experiment 1 using a procedural variation that was designed to boost control group performance. This was accomplished by reducing the number of slides in the original event and by increasing exposure duration.

Another innovation in Experiment 2 was motivated by a desire to maximize the misinformation effect. To ac-

complish this, we presented the misinformation in three postevent narratives rather than in a single narrative.

Method

The procedure used in Experiment 2 was identical to that used in Experiment 1, with a few minor modifications to be described when appropriate.

Subjects. The subjects were 338 students from the University of Washington who received class credit for their participation. They were tested in groups ranging in size from 2 to 9 persons. The standard test was given to 162 subjects, and the modified test was given to 176 subjects.

Stimuli. The slide sequence and postevent narratives were adapted from McCloskey and Zaragoza's (1985) study. The event consisted of 34 slides from the series in which a maintenance man repairs a chair and then steals some items from an office. The same four critical slides were used: the coffee jar, the magazine, the soft drink, and the tool. The same counterbalancing scheme was used.

Three narratives were presented to the subjects after the event was over. They contained approximately 272-, 293-, and 167-word descriptions of the event. They accurately described details of the event, with the exception of the critical items. The three narratives varied in writing style and in which details were emphasized. For example, the first narrative emphasized the physical appearance of the maintenance man, whereas the second narrative emphasized the appearance of the secretary. For a given subject, the narrative mentioned two critical items in a misleading way and two in a neutral way. The assignment of critical items to control and misled conditions was counterbalanced across subjects. Specifically, each version of each critical item served as a control item for about half of the subjects and as a misled item for the other half. Also, for example, of the subjects who saw a hammer in the original event, half received three narratives referring to it as a "tool" (control condition) and the other half received either three narratives referring to it as a "screwdriver" or three narratives referring to it as a "wrench" (misled condition).

Procedure. The procedure used in Experiment 2 was similar to that used in Experiment 1. The subjects saw a set of 34 slides, at a rate of 7 sec/slide. After a 10-min filler activity, the subjects read three postevent narratives. As a pretense for reading the three narratives, the subjects were told that they would have to judge which of the three writing styles was most likely to maximize learning. They read the narratives from an IBM computer screen, and each narrative was printed in a different color. After reading the narratives, each subject made his/her judgment about which writing style would maximize learning.

After approximately 5 min of filler activity, the subject took the final test on the IBM computer. The details of the testing procedure of Experiment 2 were somewhat different from those of Experiment 1 and, thus, are described in more detail here. Before the test, the subject completed 10 practice questions. For example, one practice sentence read: "The color of the floor of this room is _____. Answer 1. Black 2. White." Of 21 nonpractice test items, 17 were fillers and 4 were critical. Recognition testing for each item involved two phases; a reading phase and a test phase. In the reading phase, the subjects read a sentence with a word missing (e.g., "The man slid the calculator beneath a _____ in his box"). The instructions led the subjects to believe that they were being timed while they read the sentences and that they should read them as quickly as possible without sacrificing comprehension. In fact, the time taken to read the sentence was not actually measured. This small deception was designed to discourage the subjects from using reading time to anticipate choices or to decide on their answers.

After reading each sentence, the subject pressed the space bar with the thumb of the left hand. Doing so immediately initiated

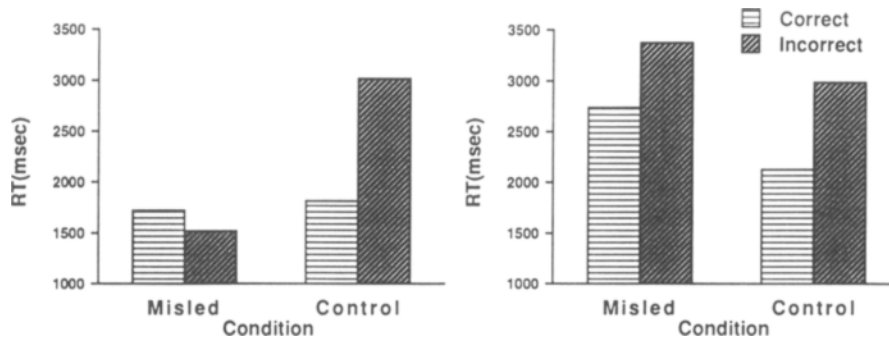


Figure 3. Mean reaction times obtained in Experiment 2. Left panel: standard test. Right panel: modified test.

presentation of the test phase and activated the timer. The subject's index finger and middle finger of the right hand rested on the 1 and 2 keys on the numeric keypad, and the subject responded by pressing one of these two buttons. In the test phase, the subjects were presented with two alternatives (e.g., "hammer," "wrench") and had to decide as quickly as possible which alternative correctly corresponded to what they had seen in the slide sequence. The subjects indicated their responses by pressing the 1 or 2 key. Instructions encouraged the subjects to respond as fast as possible without sacrificing accuracy. After every question, the subject entered his/her confidence about the accuracy of the response on a 5-point rating scale. The next question followed the confidence response by 1.5 sec. In the set of 21 test items, Questions 4, 10, 16, and 20 contained the critical items.

Results

Standard test. Once again, the subjects who took the standard test were significantly more accurate on control items than they were on misled items. Mean performance was 74% correct for control items and 43% correct for misled items, a difference of 31% [$t(161) = 8.7$, $p < .001$]. Thus, our attempt to increase control performance worked to some extent (68% in Experiment 1, 74% in Experiment 2), but not to a large degree. Moreover, our attempt to maximize the misinformation effect by presenting the postevent in three narratives did not increase the size of the misinformation event (36% impairment in performance in Experiment 1 vs. 31% in Experiment 2). Nonetheless, it is still of interest to determine

whether or not the primary RT and confidence results from Experiment 1 were replicated in Experiment 2.

Figure 3 (left panel) shows the RT data on the standard test. These data suggest that, as in Experiment 1, there was no overall tendency for the subjects to respond more slowly in the face of misinformation. In fact, they responded more quickly. Thus, there is no support for the hypothesis that misinformation produces a conflict that must be resolved during the measured RT. Furthermore, the subjects were faster when they were correct than when they were incorrect; however, on misled items, this relationship disappeared.

Statistical analyses on the RT data were carried out on 34 macrosubjects. A 2×2 within-subject analysis of variance on the macrosubject data revealed that the subjects were faster on misled than they were on control items [$F(1,33) = 39.06$, $p < .001$] and faster when they were correct than when they were incorrect [$F(1,33) = 10.55$, $p < .01$]. The interaction was significant [$F(1,33) = 34.02$, $p < .001$].

Figure 4 (left panel) shows the confidence data on the standard test. The subjects responded to misled items somewhat more, not less, confidently than they did to control items. Thus, we have no evidence to support the hypothesis that misinformation effects occur only in subjects who are simply guessing about what they saw. We also can see in the data that, on control items, the subjects were more confident about their correct answers than

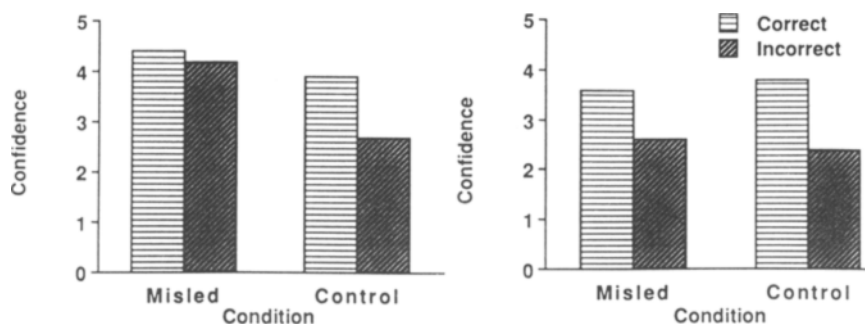


Figure 4. Mean confidence ratings obtained in Experiment 2. Left panel: standard test. Right panel: modified test.

their incorrect ones. However, on misled items, this relationship between confidence and accuracy disappeared.

Statistical analyses on the confidence data were carried out on the 34 macrosubjects. A 2×2 analysis of variance revealed that the subjects were significantly more confident when responding to misled than when responding to control items [$F(1,33) = 83.23, p < .001$]. They were more confident when they were correct than when they were incorrect [$F(1,33) = 40.33, p < .001$]. The interaction was significant [$F(1,33) = 20.06, p < .001$].

Modified test. The subjects who took the modified test were not significantly more accurate on control items than they were on misled items. Mean performance was 63% for control items and 64% for misled items [$t(175) < 1$]. This replicates the results of Experiment 1 and those of McCloskey and Zaragoza's (1985) study.

Figure 3 (right panel) shows the RT data on the modified test. The subjects responded to misled items more slowly than they did to control items. We also can see in the data that the subjects were slower when they were incorrect than when they were correct, both for control and for misled items.

Statistical analyses on the RT data were carried out on 42 macrosubjects. A 2×2 within-subjects analysis of variance on the macrosubject data revealed that the subjects responded more slowly to misled items than they did to control items [$F(1,41) = 10.17, p < .01$] and that they were faster when correct than when they were incorrect [$F(1,41) = 22.89, p < .001$]. There was no interaction ($F < 1$).

Figure 4 (right panel) shows the confidence data on the modified test. The subjects were more confident when correct than when incorrect, both for misled and correct items. An analysis of variance on the 42 macrosubjects revealed that the subjects were more confident when correct than when incorrect [$F(1,41) = 77.30, p < .001$], whereas no other main effects or interactions were significant ($F < 1$).

Discussion

The results of Experiment 2 mirrored those of Experiment 1, with only very minor variations. When the subjects were given the standard test, permitting them to choose the suggested item, they chose that item quickly and confidently. Moreover, the control subjects were more confident about their correct answers than about their incorrect ones, but this relationship between confidence and accuracy disappeared when misinformation was given. Because the pattern of results was so similar in Experiments 1 and 2, we infer that the 1.5-sec delay between the end of the sentence frame and the appearance of the response alternatives in Experiment 1 was not used for conflict resolution.

Another parallel between the two experiments can be seen in the results on the modified test. When the subjects were not permitted to choose the suggested item, they took a long time to make their choice, regardless of whether they were ultimately correct or incorrect.

GENERAL DISCUSSION

Standard Test

Consider the subjects who saw a hammer, read about a screwdriver, and then were tested with the options ("hammer" vs. "screwdriver"). In both Experiments 1 and 2, the fastest responses occurred when the subjects incorrectly selected the misinformation option ("screwdriver"). The confidence ratings on the standard test showed a similar pattern in both Experiments 1 and 2. On control items, there was a strong confidence/accuracy relationship. The subjects were more confident in their correct responses than in their incorrect ones. However, when the misinformation was presented, this relationship disappeared. Now, the subjects were just as confident about their wrong answers as they were about their right ones. Importantly, it is worth emphasizing that the misled subjects were highly confident when lured into erroneously selecting the misinformation item. That is, when they claimed they had seen the screwdriver, they were wrong but highly confident. The observation that misinformation memories are expressed with the same speed and confidence as are genuine memories should give pause to those who might have wished to use one of these characteristics as an indicator that a memory is genuine.

How would various theories of the misinformation effect explain the fast and confident selection of the misinformation option ("screwdriver")? Let us reiterate several of the major theoretical positions for why people choose the misinformation option to begin with. Some theorists have argued that subjects choose "screwdriver" because their memory for hammer has been altered (Loftus, 1989; Loftus & Hoffman, 1989; Loftus, Korf, & Schooler, 1989). However, McCloskey and Zaragoza (1985) accurately pointed out that subjects could be choosing "screwdriver" for other reasons. They could choose "screwdriver" after failing to encode both original and postevent information, but they simply guess on the test. They could choose "screwdriver" on the basis of response bias or on the basis of demand characteristics.

To understand the proposed response bias process, or what Belli (1989) refers to as *misinformation acceptance*, assume that misled and control subjects can access the original memory equally well. On those occasions when control subjects cannot remember the original item, they will guess the correct alternative half of the time. However, when misled subjects cannot remember the original item, their guesses will be biased in the direction of selecting the misinformation item on the test.

To understand the proposed demand characteristic process, or what Belli (in press) calls the *deliberation hypothesis*, assume that misled subjects remember both the original and the misinformation item. However, they respond with the misinformation item because they trust the experimenter's information more than their own memories. They perform less well than do control subjects for whom such demand characteristics are not present.

One form of the deliberation hypothesis goes as follows: At the time of the test, when subjects are asked, for example, about the critical item in the tool box, they remember both hammer and screwdriver. They mull over these choices for a while, finally settling on one of them. In some instances, they settle on the misinformation option, thus contributing to the misinformation effect. It is reasonable to presume that this group of subjects when averaged with others would increase the overall RTs. In the extreme case, if most of the misled subjects who chose the misinformation item were of this type, RTs would be expected to be quite long. They weren't. We thus suggest that this type of deliberating individual did not appear in our misled condition very often.

It is possible that we did have a significant group of slow deliberators who contributed to the misinformation effect and that this group was offset by misinformation acceptors who were fast to respond. On the basis of our data alone, we cannot rule out this possibility. It requires the assumption that misinformation acceptors (those with a memory only for the misinformation) would be relatively fast in their RTs, an assumption which is plausible but remains an open question.

In suggesting that deliberation at time of test does not play a major role in the misinformation effect that we observed here, we find support in another line of research on the misinformation effect (Belli, 1989). Belli used a "yes/no" test procedure (e.g., "Did you see a hammer?") and found that misled subjects were less likely than controls to claim to have seen the original item, but more likely than controls to correctly reject some novel item (like a wrench). Belli argued that since the yes/no test does not explicitly promote the postevent item, the deliberation hypothesis is not an appealing explanation for his effects. Its inadequacy as an explanation for the yes/no findings reduces (although does not completely eliminate) its appeal in accounting for the misinformation effect when an effect item/postevent item forced-choice test has been used. Belli goes on to argue that, since misinformation reduces the number of responses based on event information, this is likely to be, in large part, the result of misinformation interfering with the subject's ability to remember the original event.

Returning to our findings, we suggest that the fast and confident manner in which subjects embrace the misinformation option is consistent with a form of the response bias or misinformation acceptance interpretation. Misinformation could shift those who would otherwise be simply guessing, leading them to systematically choose the misleading item (as McCloskey & Zaragoza, 1985, have consistently argued). Put another way, some subjects have no memory of the critical item (e.g., tool) to begin with. Misinformation inserts a new fact into their memories, and, since they have no reason to distrust the misinformation, they become quite confident they are choosing correctly. In the process, subjects develop confidence that they actually experienced this critical fact. Such a process, if it occurs, would be thoroughly interesting and worthy

of study, despite the fact that no original memory for the critical item had ever been encoded.

Finally, it is conceivable that the original memory was altered in the process of absorbing the misinformation. A hammer was originally encoded, but it was impaired in some way when screwdriver was processed. Our results can neither confirm nor refute this possibility. When "screwdriver" is selected on the test, it is selected so quickly and confidently that, we would suggest, if there lies beneath it a genuine original pristine memory of a hammer, the trace for the hammer must be relatively impoverished. Put another way, if there exists a buried original memory, waiting to be kissed awake like Sleeping Beauty, it seems to have little effect on the behavior of subjects; despite it, they emit the fast and confident responses that are associated with cases of real memory in non-misled subjects. It is this behavior that leads us to suggest that the erroneous misinformation-induced memories are about as real to subjects as they could possibly be. In our experiments, the subjects' misinformation memories were certainly as real—in the sense of speed and confidence—as their genuine memories were.

We raised the issue of *real memory* since it has been raised by others. For example, Zaragoza and Koshmider (1989) asked whether subjects really do believe in the misinformation items that they select on a test, or whether they report the misinformation item because they believe it is an accurate depiction of what happened, not because they really remember seeing it. To drive this point home, Zaragoza and Koshmider asked readers to consider a hypothetical situation in which a witness does not remember a stop sign that was part of a witnessed event but does remember reading a report that refers to a yield sign. The witness might testify in court about seeing a yield sign simply because he/she believes the yield sign was there and wants to appear credible. That testimony might be given with confidence, whether or not the witness actually remembers seeing the sign. Similar processes could be occurring in the misinformation studies.

Thus, some investigators will probably argue that fast, confident responses do not indicate a genuine belief that the misinformation detail occurred. We, on the other hand, would argue that fast, confident responses are at least an indication of a belief in the misinformation. How else do we know when people truly believe that they experienced what they are reporting? Undoubtedly, any characteristic that one might require for a real memory could be disputed by some theorist. The category of "indicators of true BELIEF" is probably about as fuzzy as they come.

Modified Test

One clear result from both experiments concerns performance on the modified test, in which the misinformation item could not be chosen. Consider the subjects who saw a hammer, read about a screwdriver, but then were tested on "hammer" versus "wrench." These subjects were correct in their selection of "hammer" just as often

as were the control subjects. However, misled subjects responded more slowly regardless of whether they were ultimately correct in selecting "hammer" or incorrect in selecting "wrench." This pattern occurred in both Experiments 1 and 2. It is of interest to consider which theories relating to the misinformation effect would predict this pattern of results. One theory that would *not* predict this result is the rather implausible one that assumes the misinformation has no effect on subjects whatsoever. In the extreme case, if subjects had their eyes closed while the misinformation was presented, their performance would be exactly like that of control subjects. But we know, of course, that subjects do not have their eyes closed, on the basis of their behavior on the standard test. So, it must be conceded that misinformation has some influence on subjects and that the increased RTs on the modified test are a reflection of that influence. But what mechanism produces those increased RTs?

A number of plausible theories would predict longer RTs in the face of misinformation on a test that does not permit the selection of the misinformation. Longer times would be expected if misinformation leaves subjects with little or no original memory for hammer but a fairly strong memory for screwdriver. A subject who remembered a screwdriver would examine the two response options, "hammer" and "wrench," and spend some time in perplexed bewilderment before realizing that the memory must be wrong. Ultimately, "hammer" or "wrench" would be chosen, but the process would be expected to take time.

Other theorists might explain the RTs on the modified test in the face of misinformation by assuming that subjects have two memories, one for hammer and one for screwdriver. If the screwdriver memory interferes with the hammer memory, it might take subjects longer to access the hammer memory, without changing that memory at all. This thinking is reminiscent of research on the *fan* effect (Anderson, 1983, 1985; Lewis & Anderson, 1976). In the Lewis and Anderson experiment, subjects learned fantasy facts about public figures (e.g., Napoleon Bonaparte was from India). After learning these fantasy facts, subjects were tested on their memory for true facts about the public figures (e.g., Napoleon Bonaparte was an emperor). An important result was that the more fantasy facts subjects learned about an individual, like Napoleon, the longer they took to recognize a fact that they already knew about him. This is an instance in which new information can increase RTs to old information, without necessarily changing the memory for that old information.

The foregoing analysis, then, suggests two possible mechanisms by which information could increase the speed of responding on the modified test: (1) Misled subjects who have a memory for only the postevent item take longer because they must guess on the modified test; (2) misled subjects who have both event and postevent information may take longer to access the event information in memory. We believe our data suggest that both

mechanisms occur. To see this, consider that a proportion of correct responses are the result of guessing and that all of the incorrect responses are the result of guessing incorrectly. Averaging across both experiments, approximately 65% of responses were correct and 35% were incorrect. Put another way, about 35% of both misled and control subjects guessed incorrectly. In both experiments, the incorrect guesses took longer when the subjects were misled. Thus, it appears that misinformation does indeed affect the speed of pure guessing.

Does misinformation also affect the speed of accessing an original memory? Consider the 65% who were correct on the modified test. This 65% figure contains a substantial number of people who presumably had an original memory (30% with original memory and 35% guessing). Yet the increase in correct RT, from the subjects given misinformation, was just as great as the increase in RT associated with incorrect guessing. This leads us to the inference that misinformation not only increases the time it takes to guess, but it also increases the time it takes to access an original memory.

In sum, we found, as did McCloskey and Zaragoza (1985), that if one looks only at the accuracy of responding on the modified test, one is left with the impression that misinformation has no effect on performance whatsoever. In both of our experiments, the misled subjects chose the original ("hammer") response as often as did the controls. But despite this apparent similarity in performance, when one looks at speed of responding, a difference emerges. The misled subjects took longer. Although there are a variety of explanations for this result, it is of importance that subjects can appear to be behaving similarly according to one variable (accuracy) but not according to another (RTs). Further research is needed to determine the precise mechanism by which misinformation is slowing subjects down on a test in which the suggested item is not an option. One further challenge is that the precise mechanism specified must account for the finding that, with different materials, other researchers (e.g., Ceci et al., 1987b; Chandler, 1989) have shown an impairment in performance on a modified test, complicating any interpretation of performance on such tests.

REFERENCES

- ANDERSON, J. R. (1983). A spreading activation theory of memory. *Journal of Verbal Learning & Verbal Behavior*, *22*, 261-295.
- ANDERSON, J. R. (1985). *Cognitive psychology and its implications*. New York: W. H. Freeman.
- BELLI, R. F. (1988). Color blend retrievals: Compromise memories or deliberate compromise responses. *Memory & Cognition*, *16*, 314-326.
- BELLI, R. F. (1989). Influences of misleading postevent information: Misinformation interference and acceptance. *Journal of Experimental Psychology: General*, *118*, 72-85.
- BRAINERD, C. J., & REYNA, V. F. (1988). Memory loci of suggestibility development: Comment on Ceci, Ross, & Toglia (1987). *Journal of Experimental Psychology: General*, *117*, 197-200.
- CECI, S. J., & ROSS, D. F., & TOGLIA, M. P. (1987a). Age differences in suggestibility: Narrowing the uncertainties. In S. J. Ceci, M. P. Toglia, & D. F. Ross (Eds.), *Children's eyewitness memory* (pp. 79-91). New York: Springer.

- CECI, S. J., ROSS, D. F., & TOGLIA, M. P. (1987b). Suggestibility of children's memory: Psychological implications. *Journal of Experimental Psychology: General*, **116**, 38-49.
- CECI, S. J., TOGLIA, M. P., & ROSS, D. F. (1988). On remembering ... more or less: A trace strength interpretation of developmental differences in suggestibility. *Journal of Experimental Psychology: General*, **117**, 201-203.
- CHANDLER, C. C. (1989). Specific retroactive interference in modified recognition tests: Evidence for an unknown cause of interference. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **15**, 256-265.
- COLE, W. G., & LOFTUS, E. F. (1979). Incorporating new information into memory. *American Journal of Psychology*, **92**, 413-425.
- GEISELMAN, R. E. (1988). Improving eyewitness memory through mental reinstatement of context. In G. M. Davies & D. M. Thomson (Eds.), *Memory in context: Context in memory* (pp. 245-266). Chichester, England: Wiley.
- GIBLING, F., & DAVIES, G. (1988). Reinstatement of context following exposure to post-event information. *British Journal of Psychology*, **79**, 129-141.
- GREENE, E., FLYNN, M., & LOFTUS, E. F. (1982). Inducing resistance to misleading information. *Journal of Verbal Learning & Verbal Behavior*, **21**, 207-219.
- GUDJONSSON, G. H. (1986). The relationship between interrogative suggestibility and acquiescence: Empirical findings and the theoretical implications. *Personality & Individual Differences*, **7**, 195-199.
- HAMMERSLEY, R., & READ, J. D. (1986). What is integration? Remembering a story and remembering false implications about the story. *British Journal of Psychology*, **77**, 329-341.
- JOHNSON, M. K. (1987). Discriminating the origin of information. In T. F. Oltmanns & B. A. Maher (Eds.), *Delusional beliefs: Interdisciplinary perspectives* (pp. 34-65). New York: Wiley.
- JOHNSON, M. K., & HASHER, L. (1987). Human learning and memory. *Annual Review of Psychology*, **38**, 631-668.
- KOHNEN, G., & BROCKMANN, C. (1987). Unspecific postevent information, attribution of responsibility, and eyewitness performance. *Applied Cognitive Psychology*, **1**, 197-207.
- KROLL, N. E. A., & OGAWA, K. H. (1988). Retrieval of the ir retrievable: The effect of sequential information on response bias. In M. M. Gruneberg, P. E. Morris, & R. N. Sykes (Eds.), *Practical aspects of memory: Current research and issues* (Vol. 1, pp. 490-495). New York: Wiley.
- LEWIS, C. H., & ANDERSON, J. R. (1976). Interference with real world knowledge. *Cognitive Psychology*, **7**, 311-335.
- LINDSAY, D. S., & JOHNSON, M. K. (1987). Reality monitoring and suggestibility: Children's ability to discriminate among memories from different sources. In S. J. Ceci, M. P. Toglia, & D. F. Ross (Eds.), *Children's eyewitness memory* (pp. 92-121). New York: Springer.
- LOFTUS, E. F. (1989). Distortions in eyewitness memory from post-event information. In H. Wegener, F. Losel, & J. Haisch (Eds.), *Criminal behavior and the justice system: Psychological perspectives* (pp. 242-253). New York: Springer-Verlag.
- LOFTUS, E. F., & HOFFMAN, H. (1989). Misinformation and memory: The creation of new memories. *Journal of Experimental Psychology: General*, **118**, 100-104.
- LOFTUS, E. F., KORF, N., & SCHOOLER, J. W. (1989). Misinformation and memory. In J. Yuille (Ed.), *Credibility assessment: A theoretical and research perspective* (pp. 155-173). Dordrecht/Boston: Kluwer.
- 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 & Memory*, **4**, 19-31.
- MCCLOSKEY, M., & ZARAGOZA, M. (1985). Misleading postevent information and memory for events: Arguments and evidence against memory impairment hypotheses. *Journal of Experimental Psychology: General*, **114**, 1-16.
- MORTON, J., HAMMERSLEY, R. H., & BEKERIAN, D. A. (1985). Headed records: A model for memory and its failures. *Cognition*, **20**, 1-23.
- REGISTER, P. A., & KIHLSSTROM, J. F. (1988). Hypnosis and interrogative suggestibility. *Personality & Individual Differences*, **9**, 549-558.
- SCHOOLER, J. W., CLARK, C. A., & LOFTUS, E. F. (1988). Knowing when memory is real. In M. Gruneberg, P. Morris, & R. Sykes (Eds.), *Practical aspects of memory: Current research and issues* (Vol. 1, pp. 83-88). New York: Wiley.
- SCHOOLER, J. W., GERHARD, D., & LOFTUS, E. F. (1986). Qualities of the unreal. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **12**, 171-181.
- SHEEHAN, P. W., & TILDEN, J. (1986). The consistency of occurrences of memory distortion following hypnotic induction. *International Journal of Clinical & Experimental Hypnosis*, **34**, 122-137.
- 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.
- TOUSIGNANT, J. P., HALL, D., & LOFTUS, E. F. (1986). Discrepancy detection and vulnerability to misleading postevent information. *Memory & Cognition*, **14**, 329-338.
- TVERSKY, B., & TUCHIN, M. (1989). A reconciliation of the evidence on eyewitness testimony: Comments on McCloskey & Zaragoza (1985). *Journal of Experimental Psychology: General*, **118**, 86-91.
- WAGENAAR, W. A., & BOER, H. P. A. (1987). Misleading postevent information: Testing parameterized models of integration in memory. *Acta Psychologica*, **66**, 291-306.
- WELLS, G. L., & TURTLE, J. W. (1987). Eyewitness testimony research: Current knowledge and emergent controversies. *Canadian Journal of Behavioral Science*, **19**, 363-388.
- ZARAGOZA, M. S., & KOSHMIDER, J. W. (1989). Misled subjects may know more than their performance implies. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **15**, 246-255.
- ZARAGOZA, M. S., MCCLOSKEY, M., & JAMIS, M. (1987). Misleading postevent information and recall of the original event: Further evidence against the memory impairment hypothesis. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **13**, 36-44.

(Manuscript received September 15, 1988;
revision accepted for publication January 18, 1989.)