

COMMENTS

A Reconciliation of the Evidence on Eyewitness Testimony: Comments on McCloskey and Zaragoza

Barbara Tversky and Michael Tuchin
Stanford University

Loftus and her colleagues demonstrated distorted reports of memory for an event in the direction of postevent misleading information. McCloskey and Zaragoza argued that these results do not necessarily imply a weakening of the memory for the original event. They obtained evidence supporting their position by using a modified recognition test. In this experiment we introduced still another modification to the recognition test ("Yes"/"No" instead of forced choice) to answer McCloskey and Zaragoza's objections to Loftus and her colleagues' procedures. We obtained evidence for distorted reports of original information as a consequence of the misleading information. Memory and confidence data support an interference or inaccessibility interpretation of the memory errors but cannot rule out overwriting of the original information in some cases.

In a series of studies that can by now be considered classic, Loftus and her colleagues (e.g., Loftus, 1979; Loftus, Miller, & Burns, 1978) showed that misleading postevent information impairs memory report of an original event. The fate of the memory for the original event has become a matter of debate; some say that it is overwritten, others say that it is rendered less accessible, and still others say that there is no evidence that anything has happened to it at all.

In all of the experiments, subjects first viewed a sequence of slides depicting an event, such as a car accident or a theft. Next, subjects either were asked some questions about the event or read a narrative describing it. In either case, information contrary to what subjects had viewed was planted in the postevent information for misled subjects, but not for control subjects. Subjects were misled on only one or two critical items. Last, memory for the slide sequence was tested with a forced-choice recognition test. Typically, misled subjects performed more poorly than did control subjects on the critical items.

To evaluate these claims, we need to look carefully at the methods used. For the sake of simplicity, we describe the design used by Loftus et al. (1978) with the stimuli used by McCloskey and Zaragoza (1985; see Table 1, top half). Subjects watched a slide sequence of a theft in an office by a repairman presumably called in to fix a chair. A soft-drink can of, say, Coke was on one of the desks in one slide; this was one of the critical slides. After viewing the slides, subjects

answered a series of questions. For the misled subjects, one of the questions implied that there was a different soft-drink can, of, say, 7-Up. At some later time, subjects were asked which kind of soda was on the desk, Coke or 7-Up. They were told to respond on the basis of what they saw, and few, if any, reported any inconsistency between the slide sequence and the subsequent narrative. The consistent finding was that recognition memory of misled subject was lower than that of control subjects. This finding lends support to Loftus et al.'s contention that the subsequent misleading information has replaced the original information.

Even control subjects did not perform perfectly on this task; in fact, their performance hovered around 70%. McCloskey and Zaragoza (1985) argued that there is no need to assume that the misleading information has any effect on memory for the original event. Instead, McCloskey and Zaragoza made a strong case that Loftus et al.'s (1978) results would be obtained if some subjects remember the original information, some subjects remember the misleading information, and still other subjects remember nothing and are guessing. They went on to perform a series of experiments, all of which demonstrated no differences between misled and control subjects in a modified forced-choice memory task. In their modified procedure, subjects were asked to choose between the original (say, Coke) and a brand new brand of soda (say, Sunkist). McCloskey and Zaragoza argued that if memory for the original information, the Coke can, is weakened by the misleading information, then misled subjects would be more likely than control subjects to choose Sunkist over Coke. If not, then memory for Coke has not been weakened. In fact, control and misled subjects did not differ when Coke was pitted against Sunkist, which led McCloskey and Zaragoza to conclude that memory for the original Coke was not weakened as a consequence of the misleading information, 7-Up. To strengthen their case, McCloskey and Zaragoza replicated Loftus et al.'s (1978) effects with their design; that is, control subjects outperformed misled subjects when the choice was

This research was supported by National Science Foundation Grant IST 8403273 to Stanford University.

We are grateful to Maria Zaragoza and Michael McCloskey for supplying us with their stimuli and experimental procedure and for very helpful discussion. We also thank Samuel Brooks, Dale Griffin, Julie Falmagne, and Vicki Smith for their advice and assistance.

Correspondence concerning this article should be addressed to Barbara Tversky, Department of Psychology, Building 420, Stanford University, Stanford, California 94305.

Table 1
Experimental Designs for Previous and Present Research

Condition	Slides	Narrative	Forced-choice recognition		
			Loftus	McCloskey & Zaragoza	Yes-no recognition: Present design
Control	Coke	—	Coke vs. 7-Up	Coke vs. Sunkist	Coke? 7-Up? Sunkist?
Misled	Coke	7-Up	Coke vs. 7-Up	Coke vs. Sunkist	Coke? 7-Up? Sunkist?

between Coke and 7-Up. Although McCloskey and Zaragoza used a verbal recognition test, Loftus and her colleagues used a visual one. In Loftus et al.'s experiments, the misleading information was presented as presuppositions in questions about the slide sequence, and in McCloskey and Zaragoza's experiments, the misleading information was presented in a narrative presumably describing the slide sequence. Loftus et al.'s slide sequence depicted an automobile accident, and the critical choice was between stop and yield signs. These procedural details seemed to make no difference in the findings.

There are two ways in which subjects could succeed in McCloskey and Zaragoza's (1985) revised choice, between the stimulus originally viewed and an entirely new stimulus: either by correctly rejecting the new alternative or by correctly accepting the old alternative. In other words, if subjects are relatively good at knowing what they have not seen, they do not have to remember what they have seen in order to succeed in the revised recognition task. Perhaps some subjects succeeded in the modified test because they knew they had not seen a Sunkist, even if they were not sure about Coke. Even if all the discrepancy between Loftus et al.'s (1978) and McCloskey and Zaragoza's findings cannot be explained this way, it is clear that what is needed is a way to examine memory for each of the alternatives separately. A "Yes"/"No" recognition procedure in which each of the three items (the original, the misleading, and the new) are tested separately serves that purpose. Furthermore, it is possible that even if subjects do not remember items on which they are misled any differently from control items, their confidence in their memory for items on which they are misled may be different from their confidence in memory for control items.

Our experiment was designed to examine memory for each of the alternatives (original, misleading, and new) separately. It was a replication of McCloskey and Zaragoza's (1985) experiments with two important changes: Instead of forced-choice questions, questions were "Yes"/"No" (or true/false), and subjects indicated their confidence in each of their responses after answering the questions. Confidence ratings were included because we thought that if the effect of misleading information was not apparent in memory performance, it might nevertheless be detected in the more sensitive measure of confidence. The design is outlined in the bottom half of Table 1.

Method

Subjects

Seventy-two Stanford University undergraduates, volunteers or students fulfilling a course requirement, participated in the experi-

ment. They were tested in groups of 3–10 people, in a session including other brief experiments as fillers. At some point during the extensive debriefing, 3 subjects claimed to notice discrepancies; these subjects were replaced. Only 1 of those 3 noticed a discrepancy between the slide sequence and the narrative description of the event.

Stimuli and Design

The slide sequence was the same 78-slide sequence used by McCloskey and Zaragoza (1985). It depicted an office theft of a \$20 bill and a hand calculator by a maintenance man called in to repair a chair. The office had several desks and tables, cluttered with many things, some of them name-brand objects. In the sequence were two critical slides: one depicted one of three kinds of soft drinks on one of the desks (Coke, 7-Up, and Sunkist) and the other depicted one of three magazines on a table (*Mademoiselle*, *Vogue*, and *Glamour*). Each of the possibilities was presented to one third of the subjects.¹

The postevent narrative was a 694-word detailed description of the incident portrayed in the slides (McCloskey & Zaragoza's 1985 original narrative was slightly revised). The narrative mentioned the locations of many of the objects in the scene, as well as describing the action. The narratives were matched to the subjects so that each subject was misled on one of the critical items and served as a control on the other. Each of the other possibilities was used equally often as misleading information for each of the original items. The misleading sentence for the soda can was "As he was doing so, he caught sight of a set of keys on the far side of the desk near a 7-Up can." The control sentence was identical except that "soft-drink" was substituted for "7-Up." Similarly, "magazine" appeared in the control narrative instead of a proper name.

The memory test was a series of 32 true/false statements adapted from McCloskey and Zaragoza's (1985) fill-in-the-blank statements. Subjects were instructed to answer the questions on the basis of the slide sequence that they had viewed. Following each statement was a confidence scale ranging from 1 (*totally confident*) to 5 (*not at all confident*; the confidence values are reversed in the Results section). Twenty-six of the statements were fillers; the remaining six included one statement for each of the possible critical items. These were worded slightly differently and scattered through the questionnaire ("Below the magazine rack there was a copy of *Vogue* magazine," "On the table was a copy of *Glamour* magazine," and "There was a copy of *Mademoiselle* magazine on the table").

The debriefing questionnaire consisted of five questions about the experiment, successively more pointed, designed to determine whether any subjects were in any way aware of any discrepancies between the slide sequence and the narrative.

¹ Two of McCloskey and Zaragoza's (1985) critical stimuli were not used: the instant coffee jar, because it was not legible, and the tool, because it was occluded by the man's hand and because the semantics of the situation strongly suggested that only one tool was possible.

Procedure

In accordance with McCloskey and Zaragoza's (1985) procedures, subjects were told that the experiment concerned their judgments about whether the verbal or visual mode of presentation was better for memory. They were told to pay close attention to both the slides and the narrative. After receiving instructions and signing consent forms, subjects (a) viewed the slide sequence at a 5-s rate, (b) performed an unrelated 7-min filler task, (c) read the narrative once, self-paced, (d) performed a second unrelated 7-min filler task, (e) answered two questions about their intuitions on memory and mode of presentation, (f) completed the 32-item true/false recognition test based on the slides and indicated their confidence in each of their answers, and (g) filled out a debriefing questionnaire.

Results

Control subjects recognized the original stimulus significantly better than did misled subjects (sign test $z = 2.66$, $p < .005$); similarly, control subjects correctly rejected the foils far better than misled subjects rejected the misleading alternative ($z = 3.87$, $p < .001$; $z = 4.48$, $p < .001$; see Table 2). Likewise, misled subjects correctly rejected the new alternative (foil) far more successfully than they rejected the misleading alternative ($z = 3.83$, $p < .001$). Last, misled and control subjects did not differ in their rates of rejecting the new alternative(s). Thus misled subjects were in fact misled by the information presented in the postevent narrative; they accepted that alternative more than they accepted the alternative actually viewed and more than did control subjects.

Confidence data (see Table 3) corroborate these findings. For the original and new alternatives for control subjects and for the new alternative for misled subjects, confidence was significantly higher for correct answers than for errors: for control subjects, original $t(70) = 3.5$, $p < .001$, and for the foil, $t(70) = 3.8$, $p < .001$; for the misled subjects, foil, $t(70) = 2.68$, $p < .005$. These were also the cases in which overall accuracy was relatively high. The two exceptions to this pattern occurred when subjects were misled; then they were *more* confident when they incorrectly rejected the original alternative and when they incorrectly accepted the misleading alternative than when they answered correctly in these cases. The reversal was significant only for the misled alternative, $t(70) = 2.29$, $p < .05$.

Because memory for each of the alternatives was tested separately, subjects could report seeing more than one of the alternatives (see Table 4). Of the subjects in the misled condition, 19% claimed to recognize both the original and the misleading alternative; this is twice the percentage of subjects who claimed to recognize both the original and either foil in the control condition or the foil in the misled condition. Thus

Table 2
Percentage of Subjects Correct in Each Condition

Condition	Slides (Coke)	Narrative (7-Up)	Foil (Sunkist)
Control	65%	—	76%
Misled	43%	38%	76%

Table 3
Mean Confidence Ratings (1 = not at all confident, 5 = completely confident)

Condition	Slides (Coke)	Narrative (7-Up)	Foil (Sunkist)
Control			
Correct	3.8	—	3.6
Incorrect	2.7	—	2.8
Misled			
Correct	3.2	3.0	3.1
Incorrect	3.4	3.6	2.4

Note. The differences between correct confidence ratings and incorrect confidence ratings are significant for all comparisons except misled slides.

nearly half of the misled subjects who correctly recognized the original alternative also incorrectly accepted the misleading alternative.²

Discussion

To review, subjects viewed a slide sequence depicting a theft and later read a narrative describing the event that introduced one piece of misleading information. Memory for the slide sequence was tested by means of a "Yes"/"No" or true/false test that allowed separate measures of the memory for the original information, memory for the misleading information, and memory for plausible information from the same category that was entirely new. This procedure was adopted to settle a dispute concerning the effect of postevent information on memory for the original event. The results were both clear and consistent. On items on which they had been misled, subjects performed significantly and considerably worse than on items on which they were not misled in the recognition of the originally viewed soda or magazine. In fact, on items on which they were misled, subjects were more likely to incorrectly say that they had seen the item mentioned in the narrative than to correctly reject it. Moreover, misled subjects were more likely to say that they had seen the item mentioned in the narrative than the one that they had actually seen.

This constitutes substantial support for the claim that misleading information affects memory for the original information. The levels of memory are similar to those obtained by Loftus and her colleagues and by McCloskey and Zaragoza (1985) in their replications of Loftus and her colleagues'

² We are grateful to Julie Falmagne for replicating this experiment with minor changes in the filler time intervals. She obtained the same pattern of memory results: Control subjects were significantly more accurate than were misled subjects for the original information, and misled subjects incorrectly reported the misleading alternative more frequently than they reported the original alternative. She also obtained nearly the same pattern of confidence results: Control subjects' confidence was significantly higher for correct responses than for incorrect responses, yet misled subjects' confidence in errors was significantly higher than in correct responses for the misleading information.

Table 4
Percentage of Subjects Accepting or Rejecting Two Alternatives by Condition and Alternative Type

Slides	"Yes"	"No"
Control: Either foil		
"Yes"	9	56
"No"	15	20
Misled: Foil		
"Yes"	8	35
"No"	15	42
Misled: Narrative		
"Yes"	19	24
"No"	43	14

studies. Misled and control subjects were equally good at rejecting entirely new information, the new brand of soda or the magazine. This indicates that the effect of the misleading information is quite specific; it does not affect memory for the soda can or the magazine in general, but it does affect memory only for those brands seen or mentioned. The relatively high rate of correct rejections of new information also suggests why McCloskey and Zaragoza obtained no differences between misled and control subjects in their experiments. Although misled subjects may not have known what they saw, they did know what they did not see. In a forced-choice procedure, knowing what you did not see is just as good as knowing what you saw.

The confidence ratings add support to the claim that memory for the original information is affected by the misleading intervening information. In most memory experiments, subjects are more confident when they are correct than when they err. In accordance with this, subjects were, in all cases but two, significantly more confident in their correct answers than in their incorrect answers. The only reversals of that pattern were the cases in which subjects were misled. Then, subjects were more confident in their errors than in their correct responses. When misled, subjects were as certain of their errors as they were of their correct responses when not misled. In other words, when they were misled, neither subjects' memory nor their confidence discriminated correct from incorrect responses. These results argue against using confidence as a measure of likelihood of being correct, although in many applied settings, such as courtroom testimony, higher confidence is explicitly taken as an index of likelihood of truth.

Our data argue against the claim that nothing happens to the memory for the original event as a consequence of misleading information. In many subjects, memory reports are altered; moreover, subjects are quite confident about their errors. We agree with McCloskey and Zaragoza (1985) that there may be some subjects who learned only one of the soda or magazine brands and other subjects who learned nothing and were guessing. Memory performance is far from perfect.

But guessing or knowing only one brand cannot explain all our data. There were clearly some subjects who learned both brands. The "Yes"/"No" procedure allowed testing memory for each of the alternatives separately; moreover, having more than one soda can or magazine is plausible, given the scenario portrayed. It is possible, as McCloskey and Zaragoza (1985) suggested that some subjects, for reasons unrelated to the misleading information, do not remember the original information and remember only the misleading information. It has been suggested (Maria Zaragoza, personal communication, November 1987; Michael McCloskey, personal communication, November 1987) that such subjects might reject the original alternative because they remember the misleading alternative, and they believe that there could be only one can of soda or only one magazine. Although this may account for some of our effects, several lines of reasoning argue against this as accounting for all of it.

First, given the clutter of objects in the office, it was quite plausible that more than one can of soda and more than one magazine appeared in the scene. Next, it seems implausible that many subjects go through this sort of complicated counter-factual reasoning in a lengthy true/false memory test, given that they do not do that sort of reasoning in reasoning tasks (Johnson-Laird & Wason, 1978). In any case, if it is plausible that some subjects remember only the misleading information and therefore reject the original alternative, it is just as plausible that other subjects remember only the original information and therefore reject the misleading alternative. The overall effect of this sort of reasoning would be to reduce the number of subjects accepting more than one alternative but not to bias accepting one alternative over another. After we had completed our research, we learned of an experiment by Belli (1989), who also used a "Yes"/"No" recognition procedure in a McCloskey-Zaragoza task. Belli asked only one question of each item, misleading or control, and so his procedure is not open to this criticism. Nevertheless, he obtained essentially the same result: substantially worse memory of the original information in the misled subjects. Last, and of importance, there is direct evidence in our data that some subjects remember both the original and the misleading alternatives, apparently without seeing any contradiction in that. Twenty-nine percent of the misled subjects and 14% of the control subjects said that they had seen more than one of the alternatives; that is, they answered "Yes" to more than one alternative. In particular, 19% of the misled subjects said "Yes" to both the original alternative and the misleading alternative, but only 10% of the control subjects said "Yes" to the original and one of the foils. In other words, nearly half of the misled subjects who claimed to have seen the original also claimed to have seen the misleading alternative. Fewer than one sixth of the control subjects who claimed to have seen the original information also claimed to have seen one of the foils.

Our data, then, not only contradict the position that nothing has happened to the original information as a consequence of the misleading information, but they also support the position that both the original and misleading information may coexist. There is direct evidence that both types of

information are available, to at least some of the subjects. This analysis places the phenomenon of misleading information in eyewitness testimony squarely in the domain of traditional memory research.

This pattern of remembering, remembering more than one response to the same stimulus, is consistent with a venerable account of forgetting: interference theory (see Crowder, 1976; Klatzky, 1980). In a classic memory paradigm, subjects first learned one set of responses to a set of stimuli, termed the A-B list of paired associates. Then, subjects learned a second set of responses to the same stimuli, termed the A-C list. This is analogous to learning first one brand of soda or magazine and then another brand in the same situation. When subjects were tested for their memory of the A-B pairs, it was weakened, in comparison with that of a control group of subjects who did not learn the A-C list, just as the misled group in eyewitness testimony experiments did not remember the original information as well as the control group did. In later experiments in the interference tradition, researchers used memory techniques designed to elicit both the A and B responses (e.g., Barnes & Underwood, 1959). In that situation, some subjects indeed produced both responses, although others recalled only one or the other, and some failed to produce either response. This pattern of responding, similar to that observed in studies of the effects of misleading information, was obtained in spite of the fact that subjects earlier demonstrated learning of both responses. Many experiments done in the interference tradition have demonstrated that subsequent learning can reduce memory for earlier learning (retroactive interference) and that more sensitive memory tests can reveal the original learning; this indicates that at least some of the original material was still there, but it was interfered with or rendered inaccessible by the subsequent learning. Bekerian and Bowers (1983) and Christiaansen and Ochalek (1983) also offered inaccessibility as an explanation for the effects of misleading information on memory.

Interference theory provides only one framework for interpreting this phenomenon. It is easily accounted for by the many distributed models of memory currently under investigation (e.g., Hintzman, 1986; McClelland & Rumelhart, 1985). Still another compatible interpretation was provided by Lindsay and Johnson (1989) and Johnson and Suengas (1988), who suggested that subjects may fail to adequately monitor the different sources of information on the same topic. Indeed, in these sorts of experiments, subjects are misled into thinking that there are no important differences between the visual and verbal sources of information and hence that there is no need to keep track of what information came from what source. Moreover, ours are not the only data to cast doubt on the interpretations offered by McCloskey and Zaragoza (1985). Using McCloskey and Zaragoza's stimuli and design, Donders, Schooler, and Loftus (1987) replicated McCloskey and Zaragoza's findings, but they also found large differences in reaction time between misled and control subjects in the modified test. Chandler (1989) did several studies, using McCloskey and Zaragoza's procedures but quite different stimuli; in those studies, the misled subjects performed worse than the control subjects in the modified test. Last,

Belli (1989), in a design quite similar to ours, also found impairment of original learning as a consequence of intervening misleading information.

In what sense has memory for the original event been impaired by the subsequent information? We do not think that our results, or any others known to us, imply that the memory traces (whatever they are) established at the time of viewing the slides have been retroactively altered. Whether or not this seems plausible, it is not a claim that can be addressed, given current knowledge and tools. Nor do we believe that our results imply that people construct integrated situation models for events and that the situation models are rewritten when new facts come in. This may happen, but simpler explanations are at hand. Remembering, especially remembering complex events, may draw on a rich variety of stored information. That information may be more or less complete. Certain contexts and questions may elicit certain information, and other situations and probes may elicit other information. Under usual laboratory circumstances, failures to elicit information result in forgetting, an uncontroversial occurrence with which most of us are familiar. Under unusual circumstances, such as these, in which experimenters have deliberately provided contradictory information, retrieval failures may lead to distorted reports.

References

- Barnes, J. M., & Underwood, B. J. (1959). "Fate" of first-list associations in transfer theory. *Journal of Experimental Psychology*, *58*, 97-105.
- Bekerian, D. A., & Bowers, J. M. (1983). Eyewitness testimony: Were we misled? *Journal of Experimental Psychology: Learning, Memory and Cognition*, *9*, 139-145.
- Belli, R. F. (1989). Influences of misleading postevent information: Misinformation interference and acceptance. *Journal of Experimental Psychology: General*, *118*, 72-85.
- Chandler, C. C. (1989). Specific retroactive interference in modified recognition tests: Evidence for an unknown cause of interference. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*, 256-265.
- Christiaansen, R. E., & Ochalek, K. (1983). Editing misleading information from memory: Evidence for the coexistence of original and postevent information. *Memory and Cognition*, *11*, 73-86.
- Crowder, R. G. (1976). *Principles of learning and memory*. Hillsdale, NJ: Erlbaum.
- Donders, K., Schooler, J. W., & Loftus, E. F. (1987). *Troubles with memory*. Paper presented at the annual meeting of the Psychonomic Society, Seattle, WA.
- Hintzman, D. L. (1986). "Schema abstraction" in a multiple-trace memory model. *Psychological Review*, *93*, 411-428.
- Johnson, M. K., & Suengas, A. G. (1988). *Reality monitoring and judging eyewitness testimony*. Unpublished manuscript, Princeton University.
- Johnson-Laird, P. N., & Wason, P. C. (Eds.) (1978). *Thinking*. Cambridge, England: Cambridge University Press.
- Klatzky, R.L. (1980). *Human memory*. New York: Freeman.
- Lindsay, D. S., & Johnson, M. K. (in press). The eyewitness suggestibility effect and memory for source. *Memory & Cognition*.
- Loftus, E. F. (1979). *Eyewitness testimony*. Cambridge, MA: Harvard University Press.
- Loftus, E. F., Miller, D. G., & Burns, H.J. (1978). Semantic integra-

- tion of verbal information into a visual memory. *Journal of Experimental Psychology: Human Learning and Memory*, 13, 585-589.
- McClelland, J. L., & Rumelhart, D. E. (1985). Distributed memory and the representation of general and specific information. *Journal of Experimental Psychology: General*, 114, 159-188.
- 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.

Received January 11, 1988

Revision received July 11, 1988

Accepted August 5, 1988 ■