Bulletin of the Psychonomic Society
1991, 29 (4), 333-336

Indirect testing of eyewitness memory: The (non)effect of misinformation

CHAD DODSON
Princeton University, Princeton, New Jersey

and

DANIEL REISBERG
Reed College, Portland, Oregon

Misleading postevent information seems to alter memory for a previously witnessed event. But is memory for the original event "erased," or merely rendered inaccessible? In the present study, subjects were shown slides depicting a robbery. Afterwards, the subjects read a narrative containing misinformation about some details presented in the slides. Memory was then assessed both directly (with a conventional recognition test) and indirectly (with a lexical decision task). The misinformation did impair performance on the recognition test, but not on the lexical decision task. That is, decision latencies were shorter for words naming items seen in the original presentation, in comparison with baseline words; misinformation had no impact on this reaction time advantage.

Many researchers have shown that *misleading postevent information* (MLPI) seems to impair memory for a previous event. Subjects in MLPI procedures are given to-be-remembered material; then they are exposed to a description of the material, but the description contains misleading information. On a subsequent memory test, the subjects are likely to respond on the basis of the misinformation, even if they have been instructed to respond only on the basis of the originally presented material.

Many have interpreted results such as these as indicating that memory for the original material has been altered or updated by the misleading information. Others have argued, however, that the seeming memory impairment is attributable to some specific feature of the MLPI procedure. In consistency with the latter argument, subjects seem able to retrieve the original information from memory with a modified recognition test. Still other researchers, however, have reported data consistent with "updating," even with the use of modified tests. Thus, there remains a controversy about the status of the MLPI data, and about the alleged memory updating. (For recent discussions, see Belli, 1989; Lindsay & Johnson, 1989; Loftus & Hoffman, 1989; Tversky & Tuchin, 1989; and Zaragoza & McCloskey, 1989.)

The key finding in the MLPI paradigm is often negative—that is, subjects show no indication of remembering the originally presented material. In interpreting this result, we obviously want assurance that we have tested for memory in as sensitive a manner as possible. The present experiment was thus motivated by data indicating that "indirect" tests of memory are often more sensitive than direct measures such as recognition or recall tests (see Johnson & Hasher, 1987; Richardson-Klavehn & Bjork, 1988, for reviews). In addition, indirect memory tests seem to be immune to retroactive interference (Graf & Schacter, 1987; Sloman, Hayman, Ohta, & Tulving, 1988). Given the procedural similarities between MLPI and retroactive interference designs, we might expect that indirect tests will likewise be immune to MLPI.

In this procedure, we compared the effects of MLPI on a direct assessment of memory (a recognition test) and an indirect assessment (a lexical decision task). Subjects first viewed an event, presented as a series of projected slides. The subjects were then exposed to a postevent written narrative that ostensibly described the original materials accurately, but that deliberately erred on certain details. Memory for the original materials was then tested, with either a direct or an indirect test.

As in most MLPI procedures, our to-be-remembered material was presented pictorially, and the memory tests were presented in written form. Evidence indicates, however, that performance on indirect tests is greatly reduced if the materials are presented in one modality, but testing is done in a different modality (Graf, Shimamura, & Squire, 1985; Jacoby & Dallas, 1981; Jacoby & Witherspoon, 1982). Consequently, with to-be-remembered pictures, a verbal indirect test is likely to be an insensitive measure.

Given this concern, we adopted a procedure first used by Loftus (1979). In her procedure, subjects viewed a to-be-remembered event in a pictorial format. Next, the subjects answered a series of written questions about the materials, so that initial memory accuracy could be assessed. The subjects were then exposed to the MLPI, and, finally, memory was tested with a recognition test. This

Correspondence should be addressed to Chad Dodson, Department of Psychology, Princeton University, Princeton, NJ 08544, or to chad@clarity.princeton. edu.

procedure yielded the usual MLPI effect, despite the presence of the initial "accuracy test."

Our design followed this procedure of including an accuracy test, presented immediately after the slides. This served to transfer information initially seen pictorially into the same modality as that of the later indirect test. The questions for our accuracy test were therefore presented on the same computer screen as that used for the indirect test, in the same printed format. The accuracy test thus provided the basis for the memory to be assessed in the indirect test, and to be (potentially) interfered with by the MLPI.

Our direct memory test consisted of a standard recognition test, which was used to probe memory for the originally presented materials. Our indirect test consisted of a lexical decision procedure, in which the stimulus words included the names of objects and details presented earlier. Memory is revealed in this test in faster responses to previously viewed words, relative to words not seen before. The critical question, therefore, was whether the MLPI would erode this reaction time advantage, just as it erodes performance on direct memory measures.

METHOD

Subjects

Thirty undergraduates served as subjects.

Design

All subjects viewed a sequence of slides depicting the to-be-remembered event, then completed the accuracy test, and then read a misleading postevent narrative. Half of the subjects then completed the (indirect) lexical decision task, and then the forced-choice recognition test. The remaining subjects engaged in a filler activity for approximately 2 min, instead of receiving the indirect test, and then took the recognition test.

The test items consisted of 21 words, each naming a detail seen in the slides. Seven of these were assigned to each of the three conditions. Words assigned to the *baseline* condition named details depicted in the slides, but not used in the accuracy test, and not mentioned in the narrative. Performance in this condition revealed the baseline latency for these words (in the absence of any priming) in the lexical decision test. Words in the *original* condition named details seen in the slides, probed in the accuracy test, and also mentioned in the postevent narrative. Finally, words in the *misled* condition named items seen in the slides and probed in the accuracy test. However, these items were misidentified in the narrative (e.g., *wrench* instead of *hammer*; *table* instead of *cabinet*).

The assignment of words to conditions was counterbalanced through the creation of three groups, which differed only in the assignment of stimuli to condition. The subjects were randomly assigned to one of the three groups. Thus, all subjects were tested with words in all three categories.

For the direct test, we predicted poor performance for *misled* items relative to baseline or *original* items; the latter two conditions should have been equal to each other. This would simply replicate the standard MLPI effect. In the indirect memory test (lexical decision), we predicted relatively poor performance (long reaction times) in the baseline condition. These words had not been recently presented (although the items named had just been seen in pictorial form), and so we expected no priming for these. Words in the *original* condition should yield shorter reaction times, thanks to the recent priming in the accuracy test.[1] The critical result, though, was in the *misled* items. If MLPI had comparable impact on direct and indirect tests, then the *misled* condition should yield slow response times relative to the *original* condition. If, in contrast, MLPI had no impact on indirect memory testing, performance should be equal in the *misled* and *original* conditions.

Stimuli

All subjects saw a series of 79 slides depicting an incident in which a repair man enters an office and steals $20 and a calculator. (This is the same sequence of slides used in McCloskey & Zaragoza, 1985.) The slides were shown at a rate of 4 sec/slide, rear-projected onto a screen 4 ft from the subjects, with a projected size of 3 × 2 ft. Each of the test words named peripheral items in these slides, usually seen in only two or three frames.

The accuracy test, taken by all subjects, involved a 14-item questionnaire (7 questions about *original* items; 7 about *misled* items). Each question addressed details in the slides; in each case, the correct answer was the appropriate key word. The questions (four-alternative forced choice) were presented one at a time on a computer screen. There were three versions of this questionnaire for counterbalancing purposes; these differed only in their respective *original* and *misled* words.

The postevent narrative was a 1.5-pp. description of the incident seen in the slides. There were again three versions of this narrative (for counterbalancing). Each narrative presented misleading information for each of the words assigned to the *misled* condition. Words assigned to the *original* condition were referred to in the same manner as they had appeared in the slides; no reference was made to words in the baseline condition.

The items in the lexical decision task were presented at the center of a computer screen, in the same font and case as the text used in the accuracy test. The first 10 words presented were practice items (5 new words; 5 nonwords), randomly sequenced. The test items were presented next (21 test words; 21 nonwords), also randomly sequenced, with the constraint that each group of 6 words contain 3 nonwords and 1 word from each of the conditions. Each word remained on the screen until the subject's response. The screen was then blank for 900 msec before presentation of the next item.

In the recognition test (two-alternative forced choice), seven questions probed words in the *original* condition, seven probed words in the baseline condition, and seven probed words about which subjects had been misled. For both *original* and baseline questions, the incorrect choice provided was a plausible word that had not been seen before. For questions about *misled* words, the proffered lure named what had been seen in the original slides. There were two versions of this test, to allow counterbalancing of the left-right position of responses.

Procedure

The subjects were told that we were attempting to examine differences in memory for information presented in visual as opposed to written formats. The subjects were informed that the experiment consisted of the viewing of a sequence of slides, the completion of a questionnaire about the slides, the reading of a brief description of the event seen in the slides, and then the completion of a final questionnaire on what was seen. The lexical decision task was not mentioned.

The subjects were given the accuracy test immediately after they had viewed the slides, and they were instructed to circle the word that best answered each question. Following the completion of this test, the subjects were given the postevent narrative, and they were told to read it once at their own pace. Half of the subjects then completed the lexical decision task. The instructions for this task made no reference to memory for the slides or the narrative. The subjects were instead told that, in a separate study, a professor was examining reaction times for decisions about whether letter strings formed words or nonwords. The subjects were informed that the first 10 items were practice items, but they received no feedback about accuracy or errors.

After completion of either the lexical decision task or the filler activity, the subjects were given the two-alternative forced choice recognition test and told to choose the response that best answered each question. It was emphasized that they should answer the questions only in accordance with what they had seen in the slides.

RESULTS

As a preliminary analysis, we examined whether completing the indirect test contaminated performance on the subsequent direct test. The data from the recognition test were subjected to a 3 × 3 × 2 ANOVA, with the factors of condition (i.e., category of word), group (i.e., which words were assigned to which category), and test sequence

(whether the indirect test was present or absent), respectively. (The group factor was included in the analysis merely to verify that no differences existed because of counterbalancing.) The analysis showed no main effect of test sequence, nor interactions between sequence or any other factor. Apparently, the inclusion of the indirect memory test had no influence on the subsequent direct test, since the results remained unchanged when the indirect test was omitted.

Accuracy Test

The subjects made a substantial number of errors in the accuracy test (i.e., before the misleading narrative was presented). This is not surprising, for the questions probed relatively peripheral items in the slide sequence. The subjects were correct on 66% of the questions. (Chance performance would be 25%.) Errors were divided equally between words in the *original* condition and words in the *misled* condition (68% vs. 65%). Given these high error rates, the memory tests were analyzed for all items, and then they were analyzed separately for just the items that were correctly identified on the accuracy test. (The latter analysis included only 27 subjects; accuracy-test data for the remaining subjects were lost through a recording error.)

Recognition Test

The proportion of words correctly remembered on the direct memory test is shown in Table 1, for the baseline, *original*, and *misled* conditions. These data reveal a main effect of condition [$F(2,48) = 18.489, p < .01$], replicating the MLPI effect. Post hoc tests revealed a significant contrast between the *misled* and the *original* conditions [$F(1,24) = 9.102, p < .01$]. There was no reliable difference between the *original* and the baseline conditions [$F(1,24) = 1.28$], but there was a reliable contrast between the *misled* and the baseline conditions [$F(1,24) = 7.837$, $p < .01$].

As mentioned earlier, we also analyzed these data by looking only at the items that had been correctly answered on the accuracy test. (This applies only to the *original* and *misled* items, since baseline items were not included on the accuracy test.) Analysis of this subset of the data showed the same pattern as did the main analysis—a main effect of condition [$F(2,42) = 8.21, p < .002$], and no other significant main effects or interactions. Post hoc contrasts revealed a significant contrast between the *original* and *misled* conditions [$F(1,21) = 5.52, p < .03$]; the subject responded correctly to 93% of the *original* words, and 82% of the *misled* words.

Table 1
The Influence of Misleading Postevent Information on Direct and Indirect Memory Testing

	Condition		
Test	*Original*	*Misled*	Baseline
Percent recognized	78	63	74
Lexical decision latency (in milliseconds)	554	556	595

Indirect Lexical Decision Task

The results from the lexical decision task are also shown in Table 1. A 3 × 3 ANOVA (condition × group) revealed a main effect for condition [$F(2,24) = 4.579, p < .05$] but no other significant main effects or interactions. (As with the recognition analysis, the group factor was included merely to check on our counterbalancing.)

Post hoc tests revealed a significant contrast between the *original* and baseline conditions [$F(1,12) = 5.646$, $p < .05$]. There was also a reliable difference between the *misled* and baseline conditions [$F(1,12) = 5.448$, $p < .05$]. The critical contrast, however, between the *original* and the *misled* conditions, showed no indication of any difference [$F(1,12) = .028$].

The lexical decision results are virtually unchanged if we look only at items for which subjects had given correct responses in the accuracy test. The ANOVA revealed a main effect of condition [$F(2,22) = 6.101, p < .008$], and no other effects. Decisions for baseline words (M = 609 msec) were reliably slower than decisions for *original* words [M = 554 msec; $F(1,11) = 7.97, p < .02$] or *misled* words [M = 558 msec; $F(1,11) = 9.28$, $p < .01$]. There was no difference between *original* and *misled* words [$F(1,11) = .06$].

DISCUSSION

The present data replicate the MLPI memory effect with direct memory testing. In contrast, no MLPI effect was observed with indirect testing; in the lexical decision task, reaction times were reliably longer with the baseline words than with either the *original* or the *misled* words. Thus, a memory effect, in the form of priming, was observed. However, this priming effect was untouched by MLPI; that is, there was no difference in the lexical decision task between the *original* and the *misled* words.

This pattern adds to the body of data that reveal contrasts between the results of indirect and direct memory tests. Controversy remains, however, about whether direct and indirect tests are accessing separate *memory systems*, or whether these different forms of test instead involve different means of *accessing* a single underlying memory (cf. Jacoby & Dallas, 1981; Jacoby & Witherspoon, 1982; Roediger, Weldon, & Challis, 1989; Tulving & Schacter, 1990).

This controversy bears on how one should interpret the present data. If direct and indirect testing access a single memory system, our results imply that MLPI does not erase or alter memory. Instead, MLPI seems to make memory for the original material difficult to access. With a suitably sensitive memory measure (e.g., indirect testing), this memory can still be detected.

If, on the other hand, direct and indirect testing access separate memory systems, our results imply that claims made about MLPI are simply incomplete. That is, MLPI may influence one sort of memory (e.g., that revealed by direct testing), but not other sorts of memory (e.g., that revealed by indirect testing). In this case, claims about MLPI's effects must be framed more narrowly than they have been in the past.

REFERENCES

BELLI, R. F. (1989). Influences of misleading postevent information: Misinformation interference and acceptance. *Journal of Experimental Psychology: General*, **118**, 72-85.

GRAF, P., & SCHACTER, D. L. (1987). Selective effects of interference on implicit and explicit memory for new associations. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **13**, 45-53.

GRAF, P., SHIMAMURA, A. P., & SQUIRE, L. R. (1985). Priming across modalities and priming across category levels: Extending the domain of preserved function in amnesia. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **11**, 386-396.

JACOBY, L. L., & DALLAS, M. (1981). On the relationship between autobiographical memory and perceptual learning. *Journal of Experimental Psychology: General*, **110**, 306-340.

JACOBY, L. L., & WITHERSPOON, D. (1982). Remembering without awareness. *Canadian Journal of Psychology*, **36**, 300-324.

JOHNSON, M., & HASHER, L. (1987). Human learning and memory. *Annual Review of Psychology*, **38**, 631-668.

LINDSAY, D. S., & JOHNSON, M. K. (1989). The eyewitness suggestibility effect and memory for source. *Memory & Cognition*, **17**, 349-358.

LOFTUS, E. F. (1979). Reactions to blatantly contradictory information. *Memory & Cognition*, **7**, 368-374.

LOFTUS, E. F., & HOFFMAN, H. G. (1989). Misinformation and memory: The creation of new memories. *Journal of Experimental Psychology: General*, **118**, 100-104.

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**, 3-18.

RICHARDSON-KLAVEHN, A., & BJORK, R. (1988). Measures of memory. *Annual Review of Psychology*, **39**, 475-543.

ROEDIGER, H. L., III, WELDON, M., & CHALLIS, B. (1989). Explaining dissociations between implicit and explicit measures of retention: A processing account. In H. L. Roediger III & F. I. M. Craik (Eds.), *Varieties of memory and consciousness: Essays in honor of Endel Tulving* (pp. 3-41). Hillsdale, NJ: Erlbaum.

SLOMAN, S. A., HAYMAN, C. A. G., OHTA, N., & TULVING, E. (1988). Forgetting and interference in fragment completion. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **14**, 223-239.

TULVING, E., & SCHACTER, D. (1990). Priming and human memory systems. *Science*, **247**, 301-306.

TVERSKY, B., & TUCHIN, M. (1989). A reconciliation of the evidence on eyewitness testimony. *Journal of Experimental Psychology: General*, **118**, 86-91.

ZARAGOZA, M. S., & MCCLOSKEY, M. (1989). Misleading postevent information and the memory impairment hypothesis. *Journal of Experimental Psychology: General*, **118**, 92-99.

NOTE

1. One might argue that these words have been *doubly* primed, given their presentation in both the accuracy test and the narrative. However, the narrative was presented in a different format than the other materials were, and, as mentioned earlier, indirect memory tests are highly stimulus-specific. Thus, priming from the narrative seemed unlikely. In addition, "double-priming," had it occurred, would have led to a reaction time advantage for *original* words relative to *misled* words; this is contrary to the actual findings.

(Manuscript received February 14, 1991.)