Memory & Cognition
2001, 29 (8), 1120-1129

Intended and unintended effects of explicit warnings on eyewitness suggestibility: Evidence from source identification tests

KAREN L. CHAMBERS and MARIA S. ZARAGOZA
Kent State University, Kent, Ohio

Previous studies have shown that source identification (ID) tests reduce, and in some cases eliminate, eyewitness suggestibility errors. The present study showed that the suggestibility errors participants committed on a source ID test were further reduced when they were given the explicit postwarning that the experimenter was trying to trick them. These postwarnings reduced suggestibility to the same extent as prewarnings, and they did so for both once and repeatedly suggested items. In addition, the benefits of the pre- and postwarnings persisted when participants were retested 1 week later, but only if the suggestions had been repeated. For once-suggested items, the warning had the unintended effect of improving old/new recognition of the suggested information at retest, an effect that offset the improvements in source discrimination accuracy conferred by the warning. The advantages of using source ID tests for investigating group differences in eyewitness suggestibility are discussed.

The suggestibility of eyewitness memory has been a problem of longstanding interest to experimental psychologists and legal practitioners. Numerous studies have demonstrated the relative ease with which exposure to misleading postevent information can lead people to remember witnessing objects, actions, and even entire events that they never actually saw. Most studies of eyewitness suggestibility have employed variants of an experimental paradigm developed by Loftus, Miller, and Burns (1978), wherein participants who have witnessed an event are provided with misleading postevent information (via a postevent narrative account they read, or questions they answer) that contradicts (or supplements) some aspect of the originally witnessed event. For example, participants who witness a traffic accident at a stop sign might later be misinformed that the traffic sign was a yield sign. The empirical evidence for eyewitness suggestibility is the consistent finding that many misled participants incorrectly report the misinformation (e.g., yield sign) rather than the originally seen item (e.g., stop sign) on later tests of memory for the witnessed event.

In the 1980s, research on eyewitness suggestibility focused largely on its implications for basic questions about the nature of forgetting and the permanence of memory. For example, there was great interest in the possibility that exposure to misinformation caused irreversible changes in a person's memory representation of the witnessed event, and much research was aimed at assessing whether it was possible for participants who had been misinformed to "recover" memory for the witnessed event in its original form. One method used for this purpose was to warn misled participants that they had been misinformed. For example, in Greene, Flynn, and Loftus (1982), participants were warned either before or after the misleading postevent narrative that some of the information in the narrative "may be/have been inaccurate." The warning reduced suggestibility if it preceded the narrative, but was ineffective in reducing suggestibility if it occurred after participants had read the narrative. In a similar study that focused exclusively on postwarnings, Christiaansen and Ochalek (1983) found that a stronger warning ("a few of the details in the description were inaccurate", p. 469) was effective in reducing suggestibility only in those cases in which participants remembered an originally seen detail that was contradicted by the misinformation. When all of the data were considered, the effect of the postwarning was weak or nonexistent. Indeed, the only type of postwarning that has been shown to be effective is one that identifies the misinformation (e.g., "there is no question on this test for which the correct answer was mentioned in the story"; Lindsay, 1990, p. 1080; or the more extreme, "The narrative you have just read had one incorrect fact. In the slide sequence the woman did **NOT** have any cereal with her breakfast"; Wright, 1993). Yet, even such specific warnings do not always completely reverse misinformation's deleterious effects on memory performance (e.g., Lindsay, 1990). Thus, the warning studies produced mixed evidence in regard to the recoverability of original

This project was done to fulfill, in part, the requirements for the doctoral degree of K.L.C. The authors thank Ken Chambers, Sean Lane, Jennifer Ackil, Sarah Drivdahl, Kristie Payment, Karen Mitchell, and Jennifer Beck for support at various stages of the project. Preparation of this article was supported in part by NSF Grant SES-9819303 to M.S.Z. Correspondence concerning this article should be addressed to M. Zaragoza, Department of Psychology, Kent State University, Kent, OH 44242 (e-mail: mzaragoz@kent.edu) or K. Chambers, Saint Mary's College, Madeleva Hall, Notre Dame, IN 46556 (e-mail: kchamber@saintmarys.edu).

Copyright 2001 Psychonomic Society, Inc.

memories. A much clearer finding to emerge from this work was the striking ineffectiveness of nonspecific post-warnings (warnings that do not specify what aspects of the postevent narrative were inaccurate) in reducing suggestibility. It is the latter finding that is of particular relevance to the present study.

As research in eyewitness suggestibility has accumulated, attention has shifted from the recoverability issue toward a greater concern with characterizing the processes by which people come to develop false memories for suggested events. Today, there is considerable consensus that suggestibility phenomena are best construed as *source misattribution* errors. In this view, people come to have illusory recollections of suggested events because they confuse memories derived from a postevent source for memories derived from actually perceived events. As such, eyewitness suggestibility is viewed as an instance of the general difficulty people have in discriminating between related sources of information in memory (e.g., Ceci, 1995; Johnson, Hashtroudi, & Lindsay, 1993; Lindsay, 1994; Roediger, 1996; Schacter, 1999; Zaragoza & Lane, 1994).

One implication of the eyewitness suggestibility as source misattribution view is that the likelihood of committing a suggestibility error is not a fixed property of an underlying memory trace. Rather, suggestibility errors, like other misattribution errors, should be heavily influenced by the efficacy of the judgment processes and decision criteria a person employs in evaluating the origin of a memory (where decision criteria refers to the nature, amount, and number of different types of supporting information a person requires before they can attribute a memory to a particular source; see, e.g., Jacoby, Kelley, & Dywan, 1989; Johnson & Raye, 1998; and Schacter, Norman, & Koutstaal, 1998, for related ideas). In support of this claim, studies from a variety of domains have shown that misattribution errors are reduced when participants are given tests that lead them to use more effective decision criteria.

One manipulation that has been shown to reduce source misattribution errors is to give participants source identification (ID) tests (which require that participants consider that a test item may have come from several possible sources) rather than tests that query participants about a single source. For example, using an eyewitness suggestibility paradigm, Lindsay and Johnson (1989), Multhaup, de Leonardis, and Johnson (1999), and Zaragoza and Lane (1994, Experiment 3) have shown that participants were much less likely to misattribute suggested items to the witnessed event if they were given a source ID test (i.e., they were asked whether the suggested item was from the witnessed event, the postevent source, both, or neither) rather than simply being asked ("yes" or "no") whether they remembered the suggested item from the witnessed event. Presumably, when responding on the yes/no test (that had novel items as distractors), participants relied on less diagnostic, but highly accessible, information such as the suggested items' plausibility and familiarity (see Lindsay & Johnson, 1989). In contrast, because the source test required discrimination among highly familiar items, participants were forced to rely on more informative (but less accessible) source-specifying information and extended reasoning processes (for related findings, see, e.g., Jacoby, Kelley, Brown, & Jasechko, 1989, for evidence that source ID tests can reduce misattribution errors in the false fame paradigm, and Multhaup, 1995, for evidence that such tests can eliminate the elderly's increased susceptibility to such errors). Indeed, even rather subtle changes in the framing of source test questions can influence the types of misattribution errors participants make (Dodson & Johnson, 1993; Marsh & Hicks, 1998).

Although studies have shown that participants show lower levels of suggestibility on source ID tests than on recognition tests, it is important to note that source ID tests do not necessarily eliminate suggestibility effects. Indeed, a number of studies have obtained highly reliable suggestibility effects with source ID tests (Ackil & Zaragoza, 1995, 1998; Belli, Lindsay, Gales, & McCarthy, 1994; Mitchell & Zaragoza, 1996, 2001; Poole & Lindsay, 1995; Zaragoza & Lane, 1994; Zaragoza & Mitchell, 1996). This is because in the typical eyewitness suggestibility situation, the misleading suggestions are embedded in a postevent questionnaire/narrative that accurately describes the witnessed event. Hence, remembering that the suggested item was provided by the postevent source (and in many cases participants do) is not informative with regard to whether or not the suggested item was also witnessed. Thus, determining whether or not a suggested event was witnessed will require a more extensive search and evaluation of additional information in memory (e.g., expected antecedents or consequences) that might verify or disconfirm whether or not the suggested item is "real."

In the present study, we hypothesized that explicitly warning participants that the postevent source contained misleading suggestions would reduce suggestibility errors on a source ID test. Although source ID tests inform participants that the postevent questionnaire contained information that was not in the original event, they do not explicitly warn participants that the discrepant information has the potential of contaminating their memory for the witnessed event. We hypothesized that this awareness of the potentially misleading effects of the postevent information would increase the efficacy of the judgment and decision criteria that participants use in attributing memories to the witnessed event.

In an attempt to make the warning especially vivid and relevant to the participants, we staged an event in which an experimental confederate (posing as a fellow participant) publicly accused the experimenter of putting false details in the postevent questionnaire in an attempt to trick them. The experimenter responded by getting very flustered and admitting to the deception. We assumed that in this situation, the participants might also be motivated to "show" the experimenter that they could not be fooled, thus inducing them to adopt more effective judgment and decision criteria (see Zaragoza, Lane, Ackil, & Chambers, 1997, for results consistent with this possibility; and Hasher, Attig, & Alba, 1981, for the related finding that the salience of a

warning determines its efficacy in reducing the *knew-it-all-along effect*). We predicted that the participants in this postwarning group would be less suggestible than those in a no-warning group, who were treated identically but were not exposed to the warning. In addition, in order to further gauge of the effectiveness of the postwarning, we compared performance of the postwarning group to that of a prewarning group, who received a similar warning *prior* to answering the misleading questionnaire. Prewarnings are known to be highly effective in reducing suggestibility (see, e.g., Dodd & Bradshaw, 1980; Smith & Ellsworth, 1987; Underwood & Pezdek, 1998). In order to obtain a fairly comprehensive assessment of the extent to which warnings can lead people to accurately discriminate between witnessed and suggested events, we also manipulated the number of exposures to the misinformation (single or repeated). Because repetition is known to increase false memory for suggested events (e.g., Mitchell & Zaragoza, 1996, 2001; Zaragoza & Mitchell, 1996), we assessed whether the efficacy of the warning would generalize to repeated suggestions.

A second objective of the present study was to assess whether warned participants' ability to resist suggestion would change over time in a manner reminiscent of the well-known *sleeper effect* in studies of persuasion (e.g., Pratkanis, Greenwald, Leippe, & Baumgardner, 1988). It is well established that the ability to discriminate between related memories becomes more difficult over time (e.g., Schacter, Harbluk, & McLachlan, 1984). In fact, there is some evidence that warnings are ineffective in reducing suggestibility if participants are tested for the first time after a long retention interval (e.g., 1 month, Underwood & Pezdek, 1998; or 1 week, Zaragoza et al., 1997). However, because all of the relevant studies have manipulated retention interval between subjects (i.e., participants were tested *either* immediately or after a long delay), it is an open question whether a warned individual who has overtly resisted suggestion initially would show diminished resistance to suggestion if tested again at a later point in time.

The goal of eyewitness suggestibility studies is to assess the extent to which participants' memory for the witnessed event has been influenced by the misleading postevent information. For this reason, the measure that has been employed in all previous warning studies is a test of memory for the originally witnessed event. The present study differs from previous warning studies in that we also assessed whether participants remembered encountering the suggested information in the postevent questionnaire. As we will show later, assessing participants' memory for both the witnessed event and the postevent source has the advantage that it can provide some insight into the bases for group differences in suggestibility and/or changes in suggestibility over time.

METHOD

Participants and Design

A total of 230 Kent State University undergraduates completed the experiment in partial fulfillment of a course requirement, but the data from 5 warned participants who failed to remember the warning at the 1-week test and from 1 participant who was not naive (see below) were excluded from the analyses. The design was a 3 (prewarned, postwarned, or no warning group) × 3 (0, 1, or 2 exposures to suggestion) × 2 (immediate or delayed test) mixed factorial design, with the first variable manipulated between subjects. Group assignment was random. However, because the effects of prewarnings are known to be robust, fewer participants were assigned to the prewarning group ($N = 57$) than to the postwarning group ($N = 84$) and the no-warning group ($N = 83$).

Materials and Procedure

With the exception of the warning manipulation, the materials and procedure were very similar to those employed by Zaragoza and Mitchell (1996). All participants first viewed 5 min of a police training film depicting a home burglary by two youths and an ensuing car chase by the police. After completing a 10-min filler task, all participants completed a postevent questionnaire. For each participant, some of the questions were misleading in that they presupposed the existence of objects or events that, although plausible, were clearly not in the video. Note that in contrast to all previous warning studies, the misleading suggestions employed in this study were *supplementary* rather than contradictory (e.g., it was suggested that the thief carried a gun when he did not carry a weapon of any sort). Presumably, it is more difficult for participants to detect supplementary rather than contradictory misinformation, because in the supplementary case it is more difficult to prove that the suggestion is false (e.g., it is difficult to rule out the possibility that one failed to notice the gun).

Across the experiment, 12 misleading suggestions were used, each corresponding to a distinct scene from the video. For each participant, 4 suggestions were assigned to each of three exposure levels: zero, one, or two. Across participants, all suggestions served equally often in each exposure condition.

The 24-item questionnaire consisted of two subsets of 12 questions, with each question corresponding to 1 of the 12 scenes from the video. To implement the repeated-exposure manipulation, we questioned the participants about the 12 scenes in chronological order twice successively, each time querying the participants about different aspects of the same scenes. For each participant, suggestions assigned to the two-exposure (or repeated) condition appeared in each of the two questions about the relevant scene. The following example illustrates how we implemented the repeated misleading suggestion that the thief stole a ring in the third scene (when in fact the scene depicted the thief stealing money, but no jewelry): "3. After finding a *ring* and some money in the dresser, did the thief continue to look for more items in the drawers?" and "15. Having taken a ring and some money from the dresser, did the thief count the money before putting it into his pocket?" Suggestions assigned to the one-exposure (or single) condition appeared only in the second question about the relevant scene, with the misinformation (i.e., "a ring and") deleted from the first question, and suggestions assigned to the zero-exposure condition were deleted from both questions about the scene. Note that with the exception of the insertion/deletion of the suggested item, the questions remained identical across exposure conditions.

The participants in the three groups were treated identically with the following exceptions: (1) the prewarned and postwarned groups received warnings, and the no-warning group did not and (2) in the prewarned group, the warning occurred immediately prior to answering the misleading postevent questionnaire, and in the postwarned group, it occurred immediately following completion of the questionnaire. Specifically, in the prewarned group, a confederate, posing as a participant, raised his/her hand and in a confrontational voice said to the experimenter, "My friend said that there are some things in the questions that were not in the video. You're trying to trick us, right?" The participants in the postwarning group witnessed a similar incident, but did so immediately after completing the postevent questionnaire. In a similar manner to the other group, an experimen-

tal confederate raised his/her hand and in a confrontational manner said, "There were some things in the questions that weren't in the video. You're trying to trick us, right?" In both groups, the experimenter responded to the accusation by acting flustered and admitting to the discrepancy. She then informed the participants that "they might as well continue with the experiment."

Shortly after completing the postevent questionnaire, all participants received the same source-memory test. They were informed that they would hear a list of 32 statements read by the experimenter and that their task was to answer two questions about each: whether they remembered the item from the video and whether they remembered the item from the questions. The participants indicated their confidence in their responses to both questions on an answer sheet that contained two columns labeled "In the Video" and "In the Questions." Each column contained 7-point Likert-type scales with the response options *definitely yes*, *probably yes*, *maybe yes*, *unsure*, *maybe no*, *probably no*, and *definitely no*.

The test probes were in the form of simple sentences (e.g., *The thief had a gun*) and consisted of the 12 suggested items (4 each at zero, i.e., new, one, and two exposures) intermixed with 20 filler items from other sources (8 video only, 8 both video and questionnaire, and 4 new). Prior to the start of the source test, all participants were accurately informed that some of the questions that they had answered contained information that was not in the videotape and that the test list contained items from each of four possible source categories, thus minimizing the possibility that the participants in the no-warning group would claim to remember witnessing the suggested events merely because they assumed that the questionnaire was completely accurate.

After completing the source-memory test, the participants were dismissed and were asked to return to the laboratory in 1 week to finish the experiment. Upon returning to the laboratory, the participants were given the source test again (with the full set of instructions) but with the test probes presented in a different order.

At the end of the experiment, but prior to debriefing, all participants were given a brief questionnaire that was designed to assess (1) whether they were naive about the true purposes of the experiment, and (2) whether they had discussed the experiment with others during the 1-week retention interval. Only 1 participant indicated prior knowledge (she knew the confederate was an actor), and her data was excluded from the analyses. (No student indicated talking with others about it.) For the participants in the warned groups, the questionnaire also assessed their memory for the warning with the following question, "Do you remember another student raising their hand and saying something to the experimenter during the first session? If so, what did he/she say?" All but 5 participants indicated that the student had accused the experimenter of trying to trick them. The data from the 5 participants who failed to describe the warning incident were also excluded from the analyses.

The dependent variable of primary interest was the participants' suggestibility, which was operationalized as misattributions of the suggested items to the video (as measured by *yes* responses in the "Video" column). Such a response indicates a false memory of having witnessed (in the video) suggested items that were encountered only in the postevent questionnaire. (Note that it was possible for the participants to respond "yes" in both the "Video" column and in the "Questions" column, thus simultaneously making a misattribution error and a correct response.) We report the data collapsed across confidence levels (i.e., the sum of *maybe yes*, *probably yes*, and *definitely yes* responses), because analysis of the data broken down by confidence level yielded a virtually identical pattern of results.

RESULTS

For ease of exposition, the results are reported as proportions. For all analyses, statistical significance was at the $p < .05$ level. All post hoc analyses were Bonferroni pairwise comparisons.

We first asked whether nonspecific postwarnings reduced suggestibility errors at the immediate test. To this end, misattributions of suggested items to the video were submitted to a 3×3 mixed analysis of variance (ANOVA) with group (no, pre-, and postwarning) as a between-subjects variable and exposure (0, 1, 2) as a within-subjects variable. The relevant data are illustrated in the top half of Figure 1. Before reporting the results of primary interest, we first note that as expected, the participants were more likely to misattribute an item to the video when it had been suggested (1 or 2 exposures) than when it had never been presented (0 exposures), and two exposures to a suggestion produced a higher misattribution rate than a single exposure [$F(2,221) = 111.036, MS_e = 0.734, p < .05$]. Post hoc analyses confirmed higher misattribution rates for once-suggested items than never-suggested (0 exposure) items, and higher misattribution rates for repeated suggestions (2 exposure) than for single suggestions.

With regard to the main hypothesis, the results provide clear evidence that nonspecific postwarnings reduced suggestibility, and they did so regardless of whether or not the suggested items were repeated. Although the main effect of group was reliable [$F(2,221) = 6.357, MS_e = 2.877, p < .05$], it was qualified by a reliable group $\times$ exposure interaction [$F(4,442) = 4.3, MS_e = 0.73$]. As is illustrated in Figure 1, this was because the warnings reduced misattributions to the suggested items but did not reduce the base rate of misattribution errors to the never-presented (0 exposure) items. Post hoc analyses confirmed that although the prewarning and postwarning groups did not differ from the no-warning group in their misattributions of the 0-exposure items, both the prewarning and postwarning groups consistently made reliably fewer misattribution errors to the suggested items (both one exposure and two exposures) than the no-warning group.

Surprisingly, in contrast to previous findings (e.g., Greene et al., 1982), postwarnings were just as effective as prewarnings in reducing suggestibility: The pre- and postwarning groups' misattributions of suggested items did not differ at any exposure level. In addition, although repeated suggestions had a higher misattribution rate than the once-suggested items, the warnings reduced misattributions of both the repeatedly suggested and once-suggested items by the same amount. A follow-up ANOVA conducted on suggested items only (1 and 2 exposures) confirmed that the group $\times$ exposure interaction was not reliable ($F < 1$). Finally, although the warnings reduced misattribution errors, they did not eliminate them altogether: The pre- and postwarned groups' misattributions to both the once-suggested and repeatedly suggested items reliably exceeded the base rate of misattributions to the never-presented (0-exposure) items in every case.

The finding that the warnings did not affect *yes in video* responses to the never-presented (0-exposure) items argues against the possibility that the warnings simply induced a response bias against *yes in video* responses. Ad-

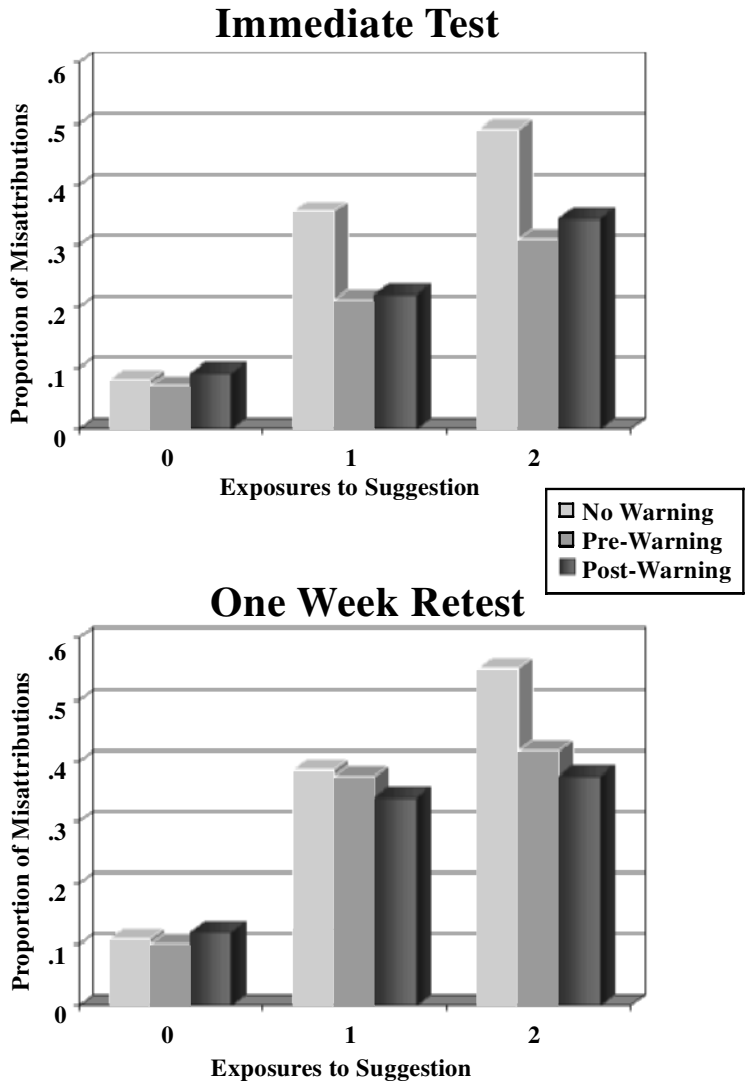

Figure 1. Mean proportion of suggested items misattributed to the video as a function of group (no warning, prewarning, and postwarning), exposure (0, 1, or 2), and retention interval (immediate or 1-week retest).

ditional evidence for this conclusion comes from the finding that the warnings did not affect the rate at which the participants responded *yes in video* to test items that they had actually seen (*M*s = 0.88, 0.89, and 0.86 for the no-, pre-, and postwarned groups, respectively), nor did the warnings affect *yes in video* responses to items that the participants had both seen in the video and read about in the postevent questionnaire (*M*s = 0.90, 0.88, and 0.87, for the no-, pre-, and postwarned groups, respectively). Hence, the warnings selectively affected performance on the suggested items.

We next asked whether the advantage conferred by the pre- and postwarnings persisted over the 1-week retention interval. Examination of the bottom half of Figure 1 shows that the answer depends on whether or not the suggested items were repeated. The data were submitted to a 3 (group: no-, pre-, and postwarning) × 3 (exposures: 0, 1, 2) × 2 (test: immediate or 1-week retest) mixed ANOVA. As can be seen by comparing the top half with the bottom half of Figure 1, the three-way interaction between group, exposure, and retention interval (immediate or 1 week) was reliable [$F(4,442) = 2.596$, $MS_e = 0.448$, $p < .05$]. At the 1-week retest, the pre- and postwarned groups no longer evidenced an advantage over the no-warning group for the once-suggested items, but both warned groups continued to evidence an advantage over the no-warning group for the repeated suggestions. Post hoc analyses conducted on the retest data confirmed that there were no group differences in misattributions of once-suggested items, but for repeatedly suggested items both the prewarning and postwarning groups committed fewer misattribution errors than the no-warning group.

In summary, we obtained a within-subjects *sleeper effect* in suggestibility (cf., Underwood & Pezdek, 1998) for the once-suggested items, but not for the repeatedly suggested items.

Having shown that pre- and postwarnings reduced misattributions of suggested items to the video in every case except the retest of the once-suggested items, we next examined the participants' memory for the suggested items in an attempt to better understand why this pattern of warning effects occurred. To this end, we first assessed the participants' ability to accurately remember having encountered the suggested items in the postevent questionnaire, as measured by the proportion of *yes in questions* responses to the suggested items. Because warnings are presumably more effective if participants can remember that the suggested items were in the questions, we were especially interested in the possibility that the sleeper effects reported above were associated with declines in memory for the once-suggested items' true source over time.

The relevant data are presented in Table 1 (misattributions of 0-exposure items to the questions were rare in all three groups and, hence, are not presented). As is obvious from the table, memory for the suggestions' true source was not related to misattribution errors in a straightforward way. A 3 (group: no-, pre-, and postwarning) × 2 (exposure: 1 or 2) × 2 (test: immediate or 1-week retest) mixed ANOVA revealed that the prewarned group had better memory for actual source than the other two groups, which in turn did not differ from each other, as is evidenced by a main effect of group [$F(2,221) = 7.5, MS_e = 1.7$] and verified by post hoc analyses. Although the prewarned group had better memory for the suggestions' true source than the postwarned group, it is interesting that the pre- and postwarned groups did not differ in their misattribution of suggested items (see Figure 1). A second finding was that repeated exposure to suggestion improved memory for actual source [i.e., the main effect of exposure was reliable; $F(2,221) = 33.6, MS_e = 1.2$]. Thus, repetition simultaneously increased misattribution errors (Figure 1) and memory for actual source (cf. Zaragoza & Mitchell, 1996). Finally, the results showed that for the once-suggested items (for which the benefits of the warning disappeared over the retest interval), there was no decline in memory for actual source across the retest interval in the warned groups. In contrast, for the repeatedly suggested items (for which the advantage of the warning persisted over the interval), there was a reliable decline in memory for actual source between tests in all groups. The test × exposure interaction was reliable [$F(2,221) = 11.6, MS_e = 0.49$], and separate ANOVAs conducted on the once-suggested and repeatedly suggested items showed a reliable effect of test for the repeatedly suggested items [$F(1,221) = 30.9, MS_e = 0.60$] but not for the once-suggested items [$F(1,221) = 1.4, MS_e = 0.71$].

Table 1
Proportion of *Yes in Questions* Responses to the Suggested Items as a Function of Group, Exposures to the Suggestion, and Test

	One Exposure		Two Exposures	
Group	Immediate Test	1-Week Retest	Immediate Test	1-Week Retest
No warning	.66	.57	.84	.71
Prewarning	.75	.79	.87	.80
Postwarning	.69	.67	.83	.71

In summary, the results show that the participants' source misattribution errors were not due to *source forgetting*. Indeed, the participants often attributed the suggestions to both the video and the questions, a finding that converges with previous studies (see, e.g., Zaragoza & Lane, 1994). Although source forgetting renders a suggested item more vulnerable to misattribution, remembering that the suggestion was provided by a postevent source does not necessarily protect the participants against misattributing the suggestion to the witnessed event.

We next assessed the extent to which the participants' remembered the content of the suggested information (sometimes referred to as *old/new recognition*). Although it is often overlooked, another factor that influences the incidence of suggestibility errors is forgetting of the suggested information. A misleading suggestion cannot affect a witness's memory unless the suggested information is preserved in memory. Put another way, in the absence of any memory for the suggested information, the participant cannot misattribute the suggestion to the witnessed event. Hence, we next assessed whether differential forgetting of the suggested information might have contributed to the pattern of warning effects we obtained.

For each group, exposure level, and test, we assessed the proportion of suggested items that the participants recognized as *old* (i.e., not forgotten) as indicated by a "yes" response to one or both of the source probes (video, questions, or both). The old/new recognition data are presented in Table 2. A 3 (group) × 2 (exposures) × 2 (test) mixed ANOVA shows that the prewarned group had higher old/new recognition than the other groups [main effect of group: $F(2,221) = 7.5, MS_e = 1.3$]. Not surprisingly, the analysis also revealed that old/new recognition was higher for repeated suggestions than for once-suggested items [main effect of exposure: $F(1,221) = 32.22, MS_e = 0.43$], and old/new recognition was higher at the immediate test than at the retest [main effect of test: $F(1,221) = 18.3, MS_e = 0.43$].

An interesting and unexpected finding was the reliable test × exposure interaction [$F(1,221) = 6.1, MS_e = 0.43$]. The left-hand side of Table 2 shows that, for the once-suggested items, the prewarned and postwarned groups did not evidence forgetting of the suggested items over the retest interval, although the no-warning group did. Post hoc analyses of the once-suggested items confirmed a reliable between-test decrease in old/new recognition for the no-warning group, but not for the pre- and postwarning groups. In contrast, post hoc analyses confirmed reliable between-test declines in old/new recognition of repeatedly suggested items for all three groups (see right-hand side of Table 2). It appears, therefore, that one unex-

Table 2
Old/New Recognition of Suggested Items as a Function of Group, Exposures to the Suggestion, and Test

	One Exposure		Two Exposures	
Group	Immediate Test	1-Week Retest	Immediate Test	1-Week Retest
No warning	.74	.65	.88	.77
Prewarning	.83	.83	.91	.82
Postwarning	.74	.73	.88	.77

pected effect of the pre- and postwarnings was that they sometimes reduced the participants' forgetting of the suggested items. To the extent that the warned participants were less likely to forget the suggested information, the warning may have had the paradoxical effect of increasing the number of misattribution errors in the warned groups. This is because the warned participants had more opportunities than unwarned participants to make these errors at the 1-week retest.

To determine whether these differences in old/new recognition contributed to the pattern of warning effects reported above, we analyzed the proportion of recognized suggestions (those identified as being from the video, questions, or both) that were misattributed to the video. As can be seen in Figure 2, this conditionalized measure of source misattributions mirrors the results of the absolute measure (see Figure 1) with one important exception: With the conditionalized measure, there was no sleeper effect in suggestibility for the once-exposed suggestions. A 3 (group) × 2 (exposure) × 2 (test) mixed ANOVA showed that the post- and prewarning groups misattributed a smaller proportion of the suggestions that they recognized than did the no-warning group, and they did so for both single and repeated suggestions and at both the immediate and 1-week retest. The main effect of group was reliable [$F(1,219) = 11.25, MS_e = 0.25, p < .05$], and group did not interact with any other variable. Post hoc analyses confirmed that the pre- and postwarning groups made fewer misattribution errors than did the no-warning group at both exposure levels and at both tests and that the pre- and postwarning groups did not differ from each other. In addition, repeated exposure to suggestion increased the proportion of recognized suggestions that were misattributed, as is evidenced by a main effect of exposure [$F(1,219) = 8.7, MS_e = 0.069$], and the proportion of recognized suggestions misattributed to the video increased over the test–retest interval [$F(1,219) = 41.55, MS_e = 0.04, p < .05$].

In summary, when groups differ in old/new recognition of target items, conditionalizing source judgments on old/new recognition provides a way of comparing source discrimination accuracy across groups. When the data were conditionalized in this way, the prewarning and postwarning groups consistently committed fewer misattribution errors than the no-warning group. Thus these results provide strong support for the conclusion that pre- and postwarnings improved source discrimination accuracy at both test intervals. However, for the once-suggested items, another, unanticipated, effect of the warnings offset the improvements in discrimination accuracy conferred by the warnings. Namely, the warnings reduced forgetting of the suggested information over the retest interval so that the warned participants were more likely than the unwarned participants to remember the suggested information at the time of the retest. Consequently, although the proportion of *recognized* suggestions that were misattributed at the retest was smaller for the warned than for the unwarned group (see Figure 2), the warned and unwarned groups made equal *numbers* of misattribution errors to the once-suggested items on the retest (see Figure 1).

DISCUSSION

The present study showed that explicitly warning participants that they had been misled reduced suggestibility errors on an immediate source ID test. In addition, postwarnings provided to the participants after they had been misled reduced their suggestibility to the same extent as prewarnings delivered prior to the misinformation, and they did so for both once- and repeatedly suggested items. It is interesting to note that even participants in the unwarned group were in some sense "warned," because the source ID test instructions informed all participants that some of the test items were from the postevent source only. Given the evidence that source ID tests reduce and, in some cases, eliminate suggestibility errors, it is striking that our explicitly warning the participants about the potentially misleading effects of the postevent questionnaire led to further reductions in suggestibility errors on a source ID test.

The present results support the hypothesis that an explicit warning can lead participants to adopt more effective decision criteria and judgment processes. The warnings did not produce a general reduction in *yes in video* test responses, but rather, selectively reduced misattributions of suggested items to the witnessed event. One possibility is that the warnings reduced misattribution errors by inducing the participants to increase the amount of evidence that they would require before attributing a memory to the witnessed event. For example, the participants in the postwarning condition may not have been willing to attribute the suggested items to the witnessed event simply on the basis that it was highly familiar, or that they were able to call up a visual image of the suggested scene (e.g., the thief stealing a ring). Rather, the warned participants may have refrained from attributing suggested items to the witnessed event unless they had other memories to support the possibility that the suggested events had actually transpired (e.g., a clear memory of what the thief did with the ring).

Why have previous studies failed to find an effect of postwarnings? The present study differed from previous warning studies in a number of ways, any of which may have increased the likelihood of obtaining a postwarning effect. For example, because all participants in the present study received a source ID test, the warned groups in our study effectively received two warnings and the un-

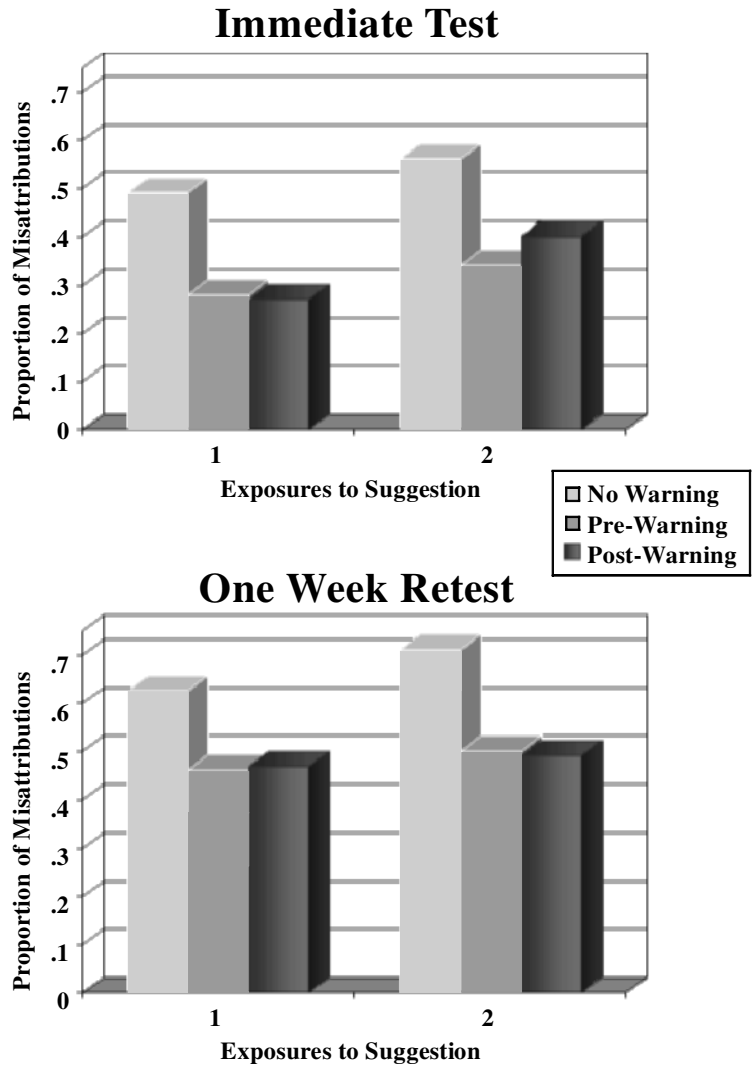

Figure 2. Misattributions of suggested items to the video conditionalized on item recognition as a function of group (no warning, prewarning, and post-warning), exposure (1 or 2), and retention interval (immediate or 1-week retest).

warned group received one. This differs from previous studies, all of which compared the effects of one warning relative to no warning. Second, the warning manipulation employed in this study (witnessing a peer angrily accuse the experimenter of trying to "trick" them) was probably more vivid, emotionally interesting, and self-relevant to the participants than the warnings used in the early studies (experimental instructions informing participants that the postevent source contained errors). The more vivid warning used here may have increased the likelihood that the participants would remember and use the warning when making source judgments on the test. Finally, the misleading suggestions employed in the present study, although plausible, included objects and events that were more consequential to the story line (e.g., *the thief had a weapon*, *the officer threatened to shoot*, *the thief stole jewelry*) than those employed in the early studies (which concerned details like the color of incidental objects). With suggestions that had greater implications for the story line, it may have been easier for the participants to use the presence/absence of supporting memories to reject the suggestions as false. The extent to which these factors, either alone or in combination, might have contributed to the warning effects reported here remains a question for future research.

Consonant with the results of previous warning studies, we also found that the postwarnings did not eliminate suggestibility. We assume that the warned participants had difficulty in discriminating between witnessed and suggested memories because there was so much overlap (or similarity) between the original event and the post-event source of information about the event (see, e.g., Johnson et al., 1993, for evidence that source overlap is a key factor in producing source confusion errors; Lind-

say & Johnson, 1989, and Mitchell & Zaragoza, 2001, for a discussion of the role of source overlap in eyewitness suggestibility effects). In addition, to the extent that the participants thought about and imagined the events (real and suggested) described in the questionnaire as they read about them, this may have rendered the suggested details confusable with "real" memories. Finally, because most of the information in the postevent questionnaire was a true account of the witnessed events, accurately remembering that the suggestions were from the postevent questionnaire did not allow the participants to infer that they were not from the witnessed event (as would be the case when the two sources are mutually exclusive). This helps to explain why the participants in this study so often misattributed the suggested items to the video, even when they could remember encountering them in the postevent questionnaire. Because this study (like most) employed misleading suggestions that were plausible and generally fit with the witnessed events, it was difficult to reject the misinformation as false on such grounds alone. Indeed, the finding that even a prewarning did not eliminate suggestibility provides further evidence of how difficult it was to identify the false details with certainty. Note, however, that the very low base rate of misattribution errors to the 0-exposure items shows that the participants did not infer the presence of the suggested items; in both the pre- and postwarned groups, suggestibility was a function of exposure to the suggestions. In summary, although our results support the conclusion that the memory-distorting effects of misinformation are to an extent "reversible," the results also show that the contaminating effects of misinformation could not be escaped altogether (cf. Wilson & Brekke, 1994).

A second question we addressed in this study was, "Do warned participants continue to make fewer suggestibility errors than unwarned participants when retested 1 week later?" As is illustrated in the bottom half of Figure 1, the answer was *yes* for repeated suggestions but *no* for the once-suggested items. Although the latter finding appears to support the conclusion that the improvements in source discrimination conferred by the warnings dissipated over the retest interval, the conditionalized analyses (bottom half of Figure 2) showed that this was not the case. On the 1-week retest, the warned participants misattributed a smaller proportion of the once-suggested items that they recognized than did the unwarned participants.

The explanation for the apparent discrepancy between Figures 1 and 2 is the finding that the warning had the unintended effect of improving the participants' long-term (i.e., 1-week) memory for the suggested information, an effect that counteracted the warning's benefits. We suspect that in an attempt to avoid being misled, the prewarned participants processed the contents of the postevent questionnaire more deeply than they normally would have, and, similarly, that the postwarned participants mentally reviewed the contents of the questionnaire more extensively than they normally would have. Whatever the mechanism responsible for this effect, the consequence was that at the time of the retest, the warned participants had more opportunities to misattribute the suggested information than their unwarned counterparts did, thus boosting the warned participants' error rate.

This unwanted effect of the warnings on the prevalence of suggestibility errors is somewhat similar to the *ironic effects* described by Wegner (1994), who has shown that attempts to avoid certain thoughts, desires, and so forth, produces effects that are diametrically opposed to the original intent. For repeatedly suggested items, the warnings did not produce this ironic effect, probably because repetition produces ironic effects of its own (see Jacoby, 1999, for other evidence of "ironic effects of repetition"). That is, because repetition alone increased the memorability of the suggested items, the additional processing that resulted from the warnings did not further increase long-term memory for the suggested information (i.e., old/new recognition of the repeatedly suggested items never differed for the warned and unwarned groups). In this regard, it is useful to note that although the warnings reduced misattribution errors to repeatedly suggested items, they did not eliminate the negative effects of repetition: Repeated suggestion increased suggestibility even in the warned groups.

The issue of primary concern in studies of eyewitness suggestibility is the assessment of the extent to which suggestive questioning results in erroneous eyewitness testimony. For this reason, eyewitness suggestibility studies have traditionally assessed participants' memory for the witnessed event, without directly assessing what participants can remember about the suggested information. The results of the present study show, however, that there are good reasons for doing the latter. Had we not directly assessed old/new recognition of the suggested information, we would not have uncovered the ironic effects of the warnings, and we might have erroneously concluded that the improvements in discrimination accuracy conferred by the warnings were in some cases very short lived. Thus, a final contribution of the present study is that it illustrates the potential advantages of using source ID tests for the investigation of eyewitness suggestibility.

REFERENCES

Ackil, J. K., & Zaragoza, M. S. (1995). Developmental differences in eyewitness suggestibility and memory for source. *Journal of Experimental Child Psychology*, **60**, 57-83.

Ackil, J. K., & Zaragoza, M. S. (1998). Memorial consequences of forced confabulation: Age differences in susceptibility to false memories. *Developmental Psychology*, **34**, 1358-1372.

Belli, R. F., Lindsay, D. S., Gales, M. S., & McCarthy, T. T. (1994). Memory impairment and source misattribution in postevent misinformation experiments with short retention intervals. *Memory & Cognition*, **22**, 40-54.

Ceci, S. J. (1995). False beliefs: Some developmental and clinical considerations. In D. L. Schacter (Ed.), *Memory distortion: How minds, brains, and societies reconstruct the past* (pp. 91-128). Cambridge, MA: Harvard University Press.

Christiaansen, R. E., & Ochalek, K. (1983). Editing misleading information from memory: Evidence for the coexistence of original and postevent information. *Memory & Cognition*, **11**, 467-475.

Dodd, D. H., & Bradshaw, J. M. (1980). Leading questions and memory: Pragmatic constraints. *Journal of Verbal Learning & Verbal Behavior*, **19**, 695-704.

Dodson, C., & Johnson, M. K. (1993). Rate of false source attributions depends on how questions are asked. *American Journal of Psychology*, **106**, 541-558.

Greene, E., Flynn, M. B., & Loftus, E. F. (1982). Inducing resistance to misleading information. *Journal of Learning & Verbal Behavior*, **21**, 207-219.

Hasher, L., Attig, M. S., & Alba, J. W. (1981). I knew it all along—or did I? *Journal of Verbal Learning & Verbal Behavior*, **20**, 86-96.

Jacoby, L. L. (1999). Ironic effects of repetition: Measuring age-related differences in memory. *Journal of Experimental Psychology: Learning, Memory, & Cognition,* **25**, 3-22.

Jacoby, L. L., Kelley, C. M., Brown, J., & Jasechko, J. (1989). Becoming famous overnight: Limits on the ability to avoid unconscious influences of the past. *Journal of Personality & Social Psychology,* **56**, 326-338.

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

Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. *Psychological Bulletin,* **114**, 3-28.

Johnson, M. K., & Raye, C. (1998). False memories and confabulation. *Trends in Cognitive Sciences*, **2**, 137-145.

Lindsay, D. S. (1990). Misleading suggestions can impair eyewitness' ability to remember event details. *Journal of Experimental Psychology: Learning, Memory, & Cognition,* **16**, 1077-1083.

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

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

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.

Marsh, R. L., & Hicks, J. L. (1998). Test formats change source-monitoring decision processes. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **24**, 1137-1152.

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

Mitchell, K. J., & Zaragoza, M. S. (2001). Contextual overlap and eyewitness suggestibility. *Memory & Cognition*, **29**, 616-626.

Multhaup, K. S. (1995). Aging, source, and decision criteria: When false fame errors do and do not occur. *Psychology & Aging*, **10**, 492-498.

Multhaup, K. S., de Leonardis, D. M., & Johnson, M. K. (1999). Source memory and eyewitness suggestibility in older adults. *Journal of General Psychology*, **126**, 74-84.

Poole, D. A., & Lindsay, D. S. (1995). Interviewing preschoolers: Effects of nonsuggestive techniques, parental coaching, and leading questions on reports of nonexperienced events. *Journal of Experimental Child Psychology*, **60**, 129-154.

Pratkanis, A. R., Greenwald, A. G., Leippe, M. R., & Baumgardner, M. H. (1988). In search of the sleeper reliable persuasion effects: III. The sleeper effect is dead. Long live the sleeper effect. *Journal of Personality & Social Psychology*, **54**, 203-218.

Roediger, H. L., III (1996). Memory illusions. *Journal of Memory & Language*, **35**, 76-100.

Schacter, D. L. (1999). The seven sins of memory: Insights from psychology and cognitive neuroscience. *American Psychologist*, **54**, 182-203.

Schacter, D. L., Harbluk, J., & McLachlan, D. (1984). Retrieval without recollection: An experimental analysis of source amnesia. *Journal of Verbal Learning & Verbal Behavior*, **23**, 593-611.

Schacter, D. L., Norman, K. A., & Koutstaal, W. (1998). The cognitive neuroscience of constructive memory. *Annual Review of Psychology*, **49**, 289-319.

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.

Underwood, J., & Pezdek, K. (1998). Memory suggestibility as an example of the sleeper effect. *Psychonomic Bulletin & Review*, **5**, 449-453.

Wegner, D. M. (1994). Ironic processes of mental control. *Psychological Review*, **101**, 34-52.

Wilson, T. D., & Brekke, N. (1994). Mental contamination and mental correction: Unwanted influences on judgments and evaluations. *Psychological Bulletin*, **116**, 117-142.

Wright, D. B. (1993). Misinformation and warnings in eyewitness testimony: A new testing procedure to differentiate explanations. *Memory*, **1**, 153-166.

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

Zaragoza, M. S., Lane, S. M., Ackil, J. K., & Chambers K. L. (1997). Confusing real and suggested memories: Source monitoring and eyewitness suggestibility. In N. L. Stein, P. A. Ornsein, B. Tversky, & C. Brainerd (Eds.), *Memory for everyday and emotional events* (pp. 401-425). Hillsdale, NJ: Erlbaum.

Zaragoza, M. S., & Mitchell, K. J. (1996). Repeated exposure to suggestions and the creation of false memories. *Psychological Science*, **7**, 294-300.

(Manuscript received June 2, 2000;
revision accepted for publication July 17, 2001.)