JOURNAL OF MEMORY AND LANGUAGE **35,** 246–260 (1996)
ARTICLE NO. 0014

Repeated Exposure to Suggestion and False Memory: The Role of Contextual Variability

KAREN J. MITCHELL AND MARIA S. ZARAGOZA

Kent State University

Although it has been well established that a single exposure to suggestion can result in the creation of false memories for suggested events, little is known about the effects of *repeated* exposure to suggestion. Zaragoza and Mitchell (in press) demonstrated that repeated exposure to postevent suggestion increased subjects' tendency to misremember witnessing the suggested information. The experiments presented here examined the possibility that increasing contextual variability between the repeated exposures would exacerbate this effect by impairing subjects' ability to discriminate accurately the precise source of the suggested items. Results from two experiments show that increasing variability by changing surface features (i.e., modality) exaggerated the deleterious effects of repeated exposure to suggestion. Increasing the spacing between exposures (Experiment 2), however, did not have the same effect.

One of the obstacles people face when they attempt to remember a specific event is discriminating between memories of the target event and other related information in memory. For example, eyewitnesses who are called to testify in a court of law must separate their memories of the event they witnessed from their own ruminations about the event as well as any pertinent information they may have acquired from other sources (e.g., newspaper accounts of the event, new information gleaned from conversations with others). Occasionally, people's attempts to distinguish between related memories fail. A potential consequence of such discrimination failures is an illusory recollection—a memory for having experienced events that were never actually experienced. It is now well established that even single exposures to misinformation can result in genuine false memories for suggested events (e.g., Ackil & Zaragoza, 1995; Belli, Lindsay, Gales, & McCarthy, 1994; Lindsay, 1990; Zaragoza & Lane, 1994; Zaragoza & Mitchell, in press). This paper, however, is concerned with illusory memories that result when people have been *repeatedly* exposed to misleading suggestions about an event they witnessed.

Interest in illusory memories is certainly not limited to the domain of eyewitness memory. For example, the possibility that people can be led to create memories of events that never happened has been attracting considerable interest of late because of its relevance to the current recovered/false memory debate. At the heart of this debate is the claim that the recent rash of alleged recovered memories of childhood abuse are in fact false memories induced by certain therapeutic practices (e.g., see Lindsay & Read, 1994; Loftus & Ketcham, 1994, for recent discussions of this claim). Like the situation in which a witness is misled, the concern here is that a suggestion introduced repeatedly by a third party might result in a false memory. In fact, it is this opportunity

These experiments composed a thesis submitted by Karen J. Mitchell in partial fulfillment of the requirements for the Master of Arts degree, Department of Psychology, Kent State University. This research was funded by National Institute of Mental Health Grant MH47858 to Maria S. Zaragoza. We thank Peggy Dombrowski, Sean Lane, John Mitchell, and Don Wolf for their patient help with stimuli preparation. We are grateful to Michelle Gilmour, Jeanine Raye, and Emily Roper for their assistance in collecting data. Address correspondence and reprint requests to Maria S. Zaragoza or Karen J. Mitchell, Department of Psychology, Kent State University, Kent, OH 44242-0001. E-mail: mzaragoz@phoenix.kent.edu. or kmitchel@kentvm.kent.edu.

0749-596X/96 $18.00

for *repeated* exposure to suggestion that is thought to make these types of situations especially conducive to false memory creation.

In spite of the recognition that repeated suggestion may encourage the induction of false memories, surprisingly little is known about repetition as a distinct factor in their development. A review of the eyewitness suggestibility literature reveals, for example, that the vast majority of these studies have assessed the effects of only a single exposure to misinformation (see Lindsay, 1994, for a review). There are a few recent studies that have employed repeated suggestion, and they report rather striking examples of false memory in which subjects claim to remember entire fictitious events, such as getting lost in a mall as a child (e.g., Ceci, Crotteau Huffman, Smith, & Loftus, 1994; Ceci, Loftus, Leichtman, & Bruck, 1994; Hyman & Billings, 1995; Hyman, Husband, & Billings, 1995; Hyman & Pentland, this issue; Loftus & Ketcham, 1994). Given the apparent facility with which false memories have been obtained in these studies, it is tempting to conclude that repetition is a powerful means of inducing false memory. However, closer examination of the methodology employed in these studies reveals that it is not at all clear to what extent repetition might have caused these effects. Because these studies were designed to mimic the complexity of real-world suggestive interview techniques, they employed procedures in which subjects were repeatedly pressed, across multiple sessions, to describe events from their childhood that never actually occurred. The locus of the interpretive difficulty is that, in all these studies, number of repeated suggestions was confounded with several other variables including the passage of time and demand (cf., Zaragoza & Mitchell, in press). Thus, the role repetition played in the creation of these memories is difficult to discern.

Warren and Lane (1995) employed a very different procedure–one that permits clearer inferences about the effects of repeated exposure to suggestion. In this study, subjects first viewed an event and then were exposed to misleading information about the event on two separate occasions. On a final test (using open-ended questions), the suggestibility of this group of subjects was compared to that of two control groups, each of which was exposed to misinformation on only one of the two occasions. Warren and Lane (1995) found that adult subjects who had been exposed to misinformation twice were no more likely to misremember witnessing the suggested information than were subjects in either of the single exposure groups. Thus, the results of this study did not support the hypothesis that repeated suggestion increases false memory.

There is, however, one study which helps establish a direct link between repeated exposure to misinformation and the formation of false memory (Zaragoza & Mitchell, in press). We describe this study in detail because it serves as the springboard for the experiments reported here. Zaragoza and Mitchell (in press) employed a procedure in which subjects viewed a videotape depicting a home burglary and later answered questions about it. Subjects were questioned about the events of the burglary in chronological order, and this process was repeated three times in succession. Each time through, subjects were asked about the same set of events, although they were asked about different aspects of these events each time. Some of the questions contained misleading suggestions. For example, when subjects were questioned about a scene in which the thief leaves the house, some subjects were asked the following question, which presupposes the misleading suggestion, ‘‘gun’’: ‘‘As the thief was leaving the house, he put his hand on the gun at his waist, looked both ways and walked out the door. Did he step out onto a porch?’’ This question was misleading because the thief in the video did not have a weapon of any sort. The critical manipulation was that for each subject some suggestions occurred once in the course of questioning and others occurred three times. So, for example, although all subjects answered questions about the thief leaving the house on three separate occasions, for some subjects the misleading suggestion ‘‘gun’’ appeared in only

TABLE 1

MEAN PROPORTION OF RECOGNIZED SUGGESTIONS ATTRIBUTED TO THE VIDEO (SOURCE MISATTRIBUTION ERRORS) AS A FUNCTION OF NUMBER OF EXPOSURES (ZARAGOZA & MITCHELL, IN PRESS)

Condition	No. of exposures 1	3	Repetition effect (3 minus 1)
Experiment 1			
Immediate Group	.21	.40	.19
48-h Delay Group	.43	.57	.14
1 Week Delay Group	.41	.59	.18
Experiment 2	.15	.28	.13

Note. Source misattribution errors were defined as the proportion of ‘‘definitely yes’’ (Experiment 1) or ‘‘remember’’ (Experiment 2) responses to the ‘‘In Video?’’ probe for recognized suggestions.

one of the three questions about the thief leaving, whereas for others ‘‘gun’’ appeared in all three questions about the scene.

In order to assess whether subjects had come to misremember witnessing the suggested information in the video, in the last phase of the experiment subjects were tested on their memory for the source of the suggested items (e.g., that the thief had a gun). That is, subjects were asked both (1) whether they remembered the suggested item from the video and (2) whether they remembered the suggested item from the questions. In Experiment 1, subjects (who were tested at one of three retention intervals) were asked to indicate their confidence in their source judgements. Experiment 2 employed a more direct measure of subjects' phenomenal experience by asking them to discriminate between items they specifically *remembered* from the video and those they *believed* were in the video (cf. Tulving, 1985). Table 1 shows that, at all retention intervals, subjects were more likely to claim that they ‘‘definitely’’ remembered seeing the suggested items in the video when they had been exposed to them three times than when they had been exposed only once. Similarly, they were more likely to claim that they ‘‘remembered’’ seeing the suggestions in the video, ruling out the possibility that repetition merely increased subjects' belief that the suggested events transpired (e.g., Arkes, Hackett, & Boehm, 1989; Bacon, 1979; Begg, Anas, & Farinacci, 1992; Begg & Armour, 1991; Hasher, Goldstein, & Toppino, 1977; Schwartz, 1982).

Why did repeated exposure to suggestion increase source misattribution errors? In attempting to answer this question, it is useful to first consider what is known about source monitoring processes in general. According to the source monitoring framework (Johnson, Hashtroudi, & Lindsay, 1993), memory for source is an attribution that results from evaluating the characteristics of the underlying memory representation. These characteristics reflect the conditions under which the memory was acquired and include contextual information (i.e., spatial/temporal information), sensory/perceptual detail, and any record of the cognitive processes engaged in during encoding. Judgements about source are made by evaluating the quantity and nature of these characteristics. Errors in source memory may occur for several reasons, including: (1) a memory has characteristics typical of another source, (2) there is an absence of characteristics that uniquely specify the item's source, and (3) the subject fails to engage in reasoning that would prevent an error (i.e., utilize general knowledge).

When subjects attempt to answer misleading questions about witnessed events they are likely to think about and imagine the information (both accurate and suggested) described in the questions (Zaragoza & Lane, 1994). At the time of test, subjects might misattribute their imagined representation of the suggested item to the video because it contains sensory/perceptual characteristics similar to memories of perceived events (Johnson et al., 1993). With repetition, the image that the subject creates of the suggested event may become increasingly elaborate, detailed, and seemingly real, increasing the likelihood of misattribution (see Suengas & Johnson, 1988, for evidence that rehearsing imagined events serves to preserve and embellish them; Zaragoza & Mitchell, in press).

Although the role of imaginal processing in the effect of repetition remains unclear, there is at least one other factor that may have contributed to this effect. Note that each repetition of the suggested information occurred in a somewhat different context (i.e., in different questions and after several intervening items) and this variability may have reduced subjects' ability to discriminate between the video and the questions as the source of the suggested information. In other words, it is likely that as a function of having encountered the suggestions in a variety of contexts, the suggested information was highly familiar yet lacking in discrete information regarding the item's source, leading subjects to overgeneralize when judging the suggested item's origin. Interestingly, William James made a similar observation about the consequences of repetition many years ago:

> If a phenomenon is met with, however, too often, and with too great a variety of contexts, although its image is retained and reproduced with correspondingly great facility, it fails to come up with any one particular setting, and the projection of it backwards to a particular past date consequently does not come about. (James, 1890/1918, p. 673)

In fact, empirical evidence that contextual variability can affect memory via the process of generalization is provided by the work of Rovee-Collier and her colleagues (see Rovee-Collier, 1991, for a review of this work). These studies utilized a contingency procedure in which infants' foot-kicking responses during training sessions moved an overhead crib mobile to provide reinforcement. Memory for the circumstances of learning were measured by the infants' response to a test mobile presented sometime later. The studies demonstrated that infants who were trained using variable training stimuli (e.g., different colored crib mobiles at each training session) were more likely to generalize the foot-kicking behavior to a novel test mobile than were infants who had been repeatedly exposed to the same stimuli (e.g., same colored crib mobile). The effects of contextual variability are also analogous to the effects of retention intervals exhibited in many paradigms. Both can produce a loss of distinct contextual/source attributes which in turn leads to generalization of responding (see Riccio, Ackil, & Burch-Vernon, 1992 for a review; cf., Rovee-Collier, 1991).

Given that contextual variability is an inherent characteristic of repetition, the goal of the present study was to manipulate that variability in order to assess whether it plays a role in the errors that result from repeated suggestion. Specifically, the present experiments were designed to investigate whether increasing the contextual variability of the repeated exposures to postevent suggestions would increase subjects' suggestibility, as measured by their tendency to make source misattribution errors.

The first experiment employed the same general three-phase procedure used by Zaragoza and Mitchell (in press). In phase 1, subjects viewed a brief video of a house burglary. During phase 2, misleading suggestions were introduced in questions about the video. Within these questions each subject received some suggestions once and others three times, while still other suggestions served as never presented control items. The innovation introduced here involved varying the context in which the thrice repeated suggestions were encountered by presenting them in different modalities. For subjects in the Single Modality Group, the three presentations of the suggestions were all in the same modality (i.e., all in print, all on videotape, or all on audiotape), a procedure analogous to that previously used by Zaragoza and Mitchell (in press). However, for subjects in the Mixed Modality Group, the suggestions were presented once in each of three different modalities (print, audiotape, and videotape), thus increasing the contextual variability of the repeated exposures. In the final phase, a source memory test was administered. The measure of primary interest was subjects' tendency to misattribute the suggested items to the originally witnessed event. It was hypothesized that increasing the variability between repeated exposures would increase false memory for the repeated suggestions, as evidenced by more source misattributions to repeated suggestions in the Mixed Modality Group than in the Single Modality Group.

EXPERIMENT 1

Method

Subjects and Design

One hundred eighty undergraduates were randomly assigned to either the Mixed Modality Group (n = 90) or the Single Modality Group (n = 90). A 2 (Mixed vs Single Modality) × 3 (0, 1, or 3 exposures) mixed design was utilized, with modality group as a between-subjects factor and number of exposures to postevent suggestion as a within subjects variable. Subjects participated in partial fulfillment of a course requirement.

Materials and Procedure

Subjects were run in small groups (i.e., ≤10). They were told that they would be participating in an experiment designed to study memory for complex events. They were further informed they would view a brief videotape and then be asked some questions about it.

Phase 1—The eyewitness event. Subjects first viewed a 5-min videotaped scene taken from a police training film. It depicted a burglary of a home by two youths and an ensuing police car chase. The clip was rich in action and dialogue.

Phase 2—Misleading postevent questioning. Immediately after seeing the video, subjects answered the same 36 postevent questions used by Zaragoza and Mitchell (in press). The set of questions was actually composed of three 12-question subsets. Each question in a subset referred to one of 12 unique events in the video (e.g., the thief entering the home). For each of these events a misleading postevent suggestion was constructed (see Appendix A). The suggestions were composed of details, action, and dialogue which supplemented, rather than contradicted, the information in the scene. The suggested items were inserted in the prefatory text of the questions and were never the focus of the question (see Table 2 for examples). The questions of each subset were ordered according to the chronology of the video. Thus, ordering was identical for each subset.

TABLE 2

EXAMPLES FROM THE POSTEVENT QUESTIONNAIRE (SUGGESTION: ''THE THIEF HAD A GUN'')

3-Exposure level

Later, as he was leaving the house the thief, putting his hand on the gun at his waist, looked both ways and went out the door. Did he slam the door behind himself?

Before leaving the house the thief checked the gun at his waist and looked both ways to see if anyone was watching. After he got out the door, did he begin to run?

As the thief was leaving the house, he put his hand on the gun at his waist, looked both ways and walked out the door. Did he step out onto a porch?

1-Exposure level

Later, as he was leaving the house the thief looked both ways and went out the door. Did he slam the door behind himself?

Before leaving the house the thief looked both ways to see if anyone was watching. After he got out the door, did he begin to run?

As the thief was leaving the house, he put his hand on the gun at his waist, looked both ways and walked out the door. Did he step out onto a porch?

Note. These questions represent the sixth question in each of the three subsets for each exposure level. For the 0-exposure level (control), there was no reference to the gun in any of the questions.

All subjects answered all 36 questions, and, by inserting suggested items in specific questions as necessary, number of exposures to suggestions was manipulated within subjects (see Table 2 for examples). For each subject, four suggestions were presented in all three subsets of questions (3-exposure level), four suggestions were presented in only the last subset of questions[1] (1-exposure level), and

[1] In a previous experiment (Zaragoza & Mitchell, in press) conducted with these same materials we showed that placement of the single-exposure suggestions (i.e., whether in the first, second, or third subset of questions) had no effect on any of the dependent measures. For this reason, to keep the number of counterbalancing groups

finally, four suggestions served as never presented control items. Counterbalancing assured that all items served at all exposure levels equally often across subjects. It is important to note that some actually perceived items were also manipulated in the questions so that they were presented either once or three times. Therefore, number of repetitions in the questions was not correlated with the accuracy of the information. These items, however, were not counterbalanced.

To allow implementation of the modality manipulation, all question subsets were individually prepared in three modalities, namely, printed, read on videotape by a male, and, read on audiotape by a female. Equal numbers of subjects in the Single Modality Group were randomly assigned to receive all 36 of their postevent questions in either printed, videotaped, or audiotaped form (*n*'s = 30 each). Subjects in the Mixed Modality Group got one 12-question subset in each of the three modalities (e.g., 12 questions in print, 12 on videotape, 12 on audiotape). Order of modalities was counterbalanced so that all modalities occurred equally often at all ordinal positions. In all then, counterbalancing assured that across subjects, all suggestions occurred equally often at all exposure levels in all modalities and that all modalities were represented equally at all ordinal positions.

Instructions varied somewhat for each group depending on how the first 12-question subset was to be administered. However, they were constructed to be equated for length and of parallel semantic structure. Depending on the group, subjects were told either that they would hear an audiotape of someone reading them questions about the video they had just seen, see and hear someone reading them questions on videotape, or, be given a sheet with questions printed on it. Subjects who were to receive the questions via videotape or audiotape were told they would have 8 s in which to answer each question. Written questions were answered at the subjects' own pace. (Time to complete the questions was approximately equal for all groups, about 5 min.) Subjects were told that they should answer each and every question even if they had to guess.

After subjects' inquiries were addressed, the first 12 questions were delivered. When all subjects were finished, instructions were given for the next subset of questions. Again, these varied depending on how the questions were to be delivered, but were similar to those given for the first subset of questions. As a ruse for the additional questions, subjects were told either: (1) we were interested in investigating the effect of a different modality (it was named) on people's recall of information about a movie (Mixed Modality) or (2) we were interested in investigating the effect of answering additional questions on people's recall of information about a movie (Single Modality). The same procedure followed for the final set of 12 questions. Note that procedure and instructions were parallel for all groups, thus keeping time lag between question subsets relatively equal.

When all 36 questions had been delivered subjects engaged in a 7-min filler task.

Phase 3—Source memory test. A surprise source memory test followed the 7-min filled interval. The probes, which were presented in the same randomized order to all subjects, were 32 statements read on a cassette recorder in a male voice, so as to be as different from all other stimuli exposures as possible. The interitem interval was 8 s. Twelve of these statements contained the critical items (e.g., ''The thief had a gun.''). For any one subject, 4 of these were novel items (i.e., 0 exposures) and 8 were suggested items that appeared in the postevent questions only (4 each at 1 and 3 exposures). The remaining 20 test probes were filler items that were included solely for purposes of ensuring that the test list contained equal numbers of test probes from each of the four possible source categories (i.e., video only, questions only, both video and questions, and neither video nor questions). Specifically, because the critical items contributed 8 probes from the ''questions only'' category and 4 probes that were ''new'' to the test list,

manageable, we embedded all single-exposure suggestions in the last set of questions.

the filler items consisted of 8 ''video only'' items, 8 ''both'' items, and 4 additional ''new'' items. It should be noted that filler items were chosen with the goal of giving subjects a clear benchmark against which to evaluate their memories of the suggested items, rather than with the goal of assessing subjects' general source monitoring ability. Therefore, these items were selected to be very obvious members of their source category. For example, we chose as ''video only'' filler items highly salient objects that had been visible throughout large segments of the video but were never mentioned in the questions.

Subjects were given both written and verbal instructions for the source memory test. They were told that they would hear 32 statements read to them on a tape recorder at 8-s intervals. All subjects were explicitly informed that of the 32 statements they would hear, some contained information that was only in the video of the burglary scene, some contained information that was not in the video of the burglary scene but was contained in the questions they answered, some contained information that was in both the video and the questions, and finally, some of the test statements contained information that was in neither the video of the burglary nor the questions. Including this explicit warning about misinformation allows one to rule out the possibility that subjects would merely believe that the misinformation was contained in the video because the experimenter said it was so. Rather they were told they were to base their source judgements on their memory of the events. They were instructed that for each statement they heard they were to answer two questions. Did they remember the information contained in the test probe: (1) from the video of a burglary they saw? and (2) from the postevent questions?

For each test probe, subjects circled ''Yes'' or ''No'' for each of these two questions on their answer sheet, which contained two columns, one labeled ''Video?'' and one labeled ''Questions?'' Care was taken to differentiate for the subjects what was meant by *video* and *questions*. It was made clear to subjects that *questions* here referred to *any* of the questions they had answered about the burglary video. To ensure that subjects' responses accurately reflected their memory for the source of the test items, the experimenter went through each of the four possible source categories (i.e., video only, questions only, both, neither), and indicated what the appropriate response pattern should be. For example, subjects were told that if they remembered the information contained in the test statement from the video only, they should circle ''yes'' in the ''Video'' column and ''no'' in the ''Questions'' column. They were cautioned that they must make two responses to each test item.

The question of primary interest in this study was whether subjects' memory for the source of the repeated suggestions varied as a function of contextual variability, and for this reason we present the results for the suggested items only.[2] A ''yes'' response to a suggested item in the ''Video'' column indicated a source misattribution error, while a ''yes'' response to a suggestion in the ''Questions'' column indicated a correct source judgement. Note that these are not mutually exclusive judgements. A subject could both misattribute a suggested item to the video and correctly attribute it to the questions.

Results and Discussion

Item recognition. A distinction can be made between subjects' ability to recognize a suggested item as ''old'' and their ability to identify the suggested item's source. We measured item recognition as subjects' ability to identify a suggestion as being from either the video and/or the questions. As can be seen in the top half of Table 3, overall, item recognition was better for suggestions presented three times than for those presented once ($F(1,178) = 52.31$, $MSE = .02$, $p < .0001$). Of greater interest is the finding that item recognition did not vary between the Single and the Mixed

[2] Note also that analysis of the filler item data is not informative given that these items were selected so that subjects could easily identify their source and performance thus tended toward ceiling.

TABLE 3

MEAN PROPORTION OF SUGGESTIONS RECOGNIZED AS A FUNCTION OF GROUP AND NUMBER OF EXPOSURES IN EXPERIMENTS 1 AND 2

Condition	No. of Exposures 1	3
Experiment 1		
Single Modality	.86	.97
Mixed Modality	.84	.96
Experiment 2		
Single Modality Consecutive	.86	.97
Single Modality Spaced	.81	.97
Mixed Modality Consecutive	.81	.94
Mixed Modality Spaced	.81	.96

Modality groups ($F(1,178) = .72$, $MSE = .03$, $p > .10$), in spite of the fact that contextual variability has previously been shown to improve item memory (see Greene, 1992, for a review). However, ceiling effects in the 3-exposure case may have precluded finding differences between the groups.

Obviously, one cannot attribute a source to an item that one does not remember. Therefore, to control for the item recognition differences between exposure conditions, the source data reported below (and throughout) were conditionalized on item recognition. That is to say, for each source (video or questions), we report the proportion of recognized suggestions that were attributed to that source.

Source misattribution errors. The measure of primary concern was the extent to which subjects misattributed the suggested items to the video.[3] Figure 1 shows that, overall, repeated exposure to suggestion increased source misattribution errors as evidenced by more errors in the 3-exposure condition than in the 1-exposure condition ($F(1,178) = 39.98$, $MSE = .05$, $p < .0001$). This result replicates the repetition effect reported by Zaragoza and Mitchell (in press). More importantly, Fig. 1 clearly shows that although the Mixed Modality Group committed more errors overall ($F(1,178) = 4.00$, $MSE = .25$, $p < .05$), the detrimental effect of repeated exposure was greater for the Mixed than the Single Modality Group, as evidenced by a significant Group X Exposures interaction ($F(1,178) = 5.82$, $MSE = .05$, $p < .05$). Planned comparisons confirmed that there was a significant repetition effect in both groups ($F(1,89) = 8.64$, $MSE = .04$, $p < .04$, for the Single Modality Group; $F(1,89) = 34.23$, $MSE = .05$, $p < .0001$, for the Mixed Modality Group) and that the locus of the interaction was a difference between the groups at the 3-exposure level ($F(1,178) = 7.58$, $MSE = .15$, $p < .01$) but not at the 1-exposure level ($F(1,178) = .85$, $MSE = .15$, $p > .10$). This finding helps rule out a global confusion account of this contextual variability effect. In other words, it does not appear that Mixed Modality subjects were more confused in general, but rather the confusion was limited to those items to which they were actually exposed in all three modalities.

[3] Base rates of false alarms (i.e., ''yes'' responses to ''Video?'' for items at the 0-exposure level) were very low (M's = .09 and .07 for the Single and Mixed Modality Groups, respectively) and did not vary between groups ($p > .05$). Nor did performance on the 0-exposure items vary on any other dependent variable in either experiment of the present study (all p's $> .05$). Therefore, since the primary interest was in determining whether the effects of repetition vary differentially as a function of contextual variability, we focus on the 3-exposure vs 1-exposure level comparison.

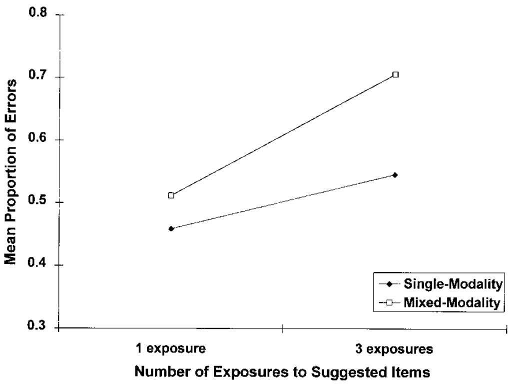

FIG. 1. Mean proportion of recognized suggestions attributed to the video (source misattribution errors) as a function of group and number of exposures in Experiment 1.

TABLE 4

MEAN PROPORTION OF RECOGNIZED SUGGESTIONS ATTRIBUTED TO THE POSTEVENT QUESTIONS (ACTUAL SOURCE MEMORY) AS A FUNCTION OF GROUP AND NUMBER OF EXPOSURES IN EXPERIMENTS 1 AND 2

	No. of Exposures	
Condition	1	3
Experiment 1		
Single Modality	.91	.99
Mixed Modality	.85	.95
Experiment 2		
Single Modality Consecutive	.93	.97
Single Modality Spaced	.89	.94
Mixed Modality Consecutive	.89	.98
Mixed Modality Spaced	.87	.97

Memory for actual source. The top half of Table 4 shows the proportion of recognized suggestions that were correctly attributed to the postevent questions. Overall, repeated exposure to suggestion improved memory for the suggested items' actual source ($F(1,178) = 34.72$, $MSE = .02$, $p < .0001$), a finding consistent with those of Zaragoza and Mitchell (in press). Of greater interest is the finding that Mixed Modality subjects made fewer correct source attributions than Single Modality subjects overall ($F(1,178) = 5.90$, $MSE = .04$, $p < .05$) (see the top half of Table 4), although the Group X Exposures interaction was not significant ($F(1,178) = .11$, $MSE = .02$, $p > .10$). The results imply that subjects who were exposed to suggestions in more varied contexts had greater difficulty remembering the actual source of all the suggested items, not just those that were repeated in varied contexts.

Taken together these results support the conclusion that contextual variability weakens subjects' ability to make accurate source attributions.

EXPERIMENT 2

Experiment 2 was designed to replicate and extend the findings of Experiment 1. To better understand how contextual variability may come to affect source judgements in this repetition paradigm, two changes were made in the methodology. First, we utilized a measure of source memory that allowed subjects to rate their confidence in their source memory and included ''unsure'' as a response option (Zaragoza & Lane, 1994, Experiment 5). Because subjects can choose ''unsure'', this measure allows one to partial out the possible contribution of guessing to performance. Second, we sought to assess whether the results of Experiment 1 would generalize to a form of contextual variation which occurs naturally in real-world repetition situations, namely, spacing of exposures. This was accomplished by introducing a 5-min filled interval between postevent question subsets. We reasoned that the filled interval would increase contextual variability, most notably because subjects returning to the postevent questions after a distractor task should be in a different frame of mind upon return than should subjects who work on the question subsets consecutively. Four groups were formed by orthogonally varying modality and spacing. In other words, there were two groups which replicated those in Experiment 1 (Single and Mixed Modality Consecutive groups) and two which represented an extension (Single and Mixed Modality Spaced groups).

Method

Subjects and Design

Subjects were 396 undergraduates who participated in partial fulfillment of a course requirement. They were randomly assigned to one of four groups, Single Modality Consecutive ($n = 108$), Mixed Modality Consecutive ($n = 108$), Single Modality Spaced ($n = 90$), or Mixed Modality Spaced ($n = 90$), as described below. This resulted in a 2 (Modality: Single vs Mixed) X 2 (Spacing: Consecutive vs Spaced) X 3 (Exposure Level: 0, 1, or 3 exposures) mixed design, with Modality and Spacing as between-subject factors and Exposure Level as a within-subject variable.

Materials and Procedure

With the exception of two bogus filler tasks (described below) and the nature of the response options on the source memory test (described below), the materials and procedure were the same as those used in Experiment 1.

Subjects were informed that they were participating in an experiment designed to investigate various cognitive abilities including judgement and memory. They were told that one task would involve watching a short videotape and answering some questions about it, but that there would be a series of different tasks included in the experimental session.

Except for the two additional filler tasks, all subjects were exposed to the same general three-phase procedure used in Experiment 1. Subjects in the two Consecutive groups performed the two filler tasks first, followed by phase 1 and phase 2, as explained in Experiment 1. For subjects in the Spaced groups one filler task was completed after each of the first two postevent question subsets.

The filler tasks were as follows:

Violence rating task. Ten 20-s segments of popular music were recorded on cassette with 8-s interitem intervals. Subjects were informed they would hear a series of 20-s music clips played on a cassette recorder. For each clip their task was to: (1) make a ‘‘yes/no’’ judgement about whether they recognized the song, and, (2) rate the violence of the content of the song on a 7-point Likert-type scale with anchors at 1 (*not at all violent*) and 7 (*very violent*). Subjects circled their responses on a printed answer sheet.

Word association task. Stimuli were 15 number–word pairs, taken from MacLeod (1988), presented individually at 20-s intervals via slide projector. The pairs were carefully selected so as not to overlap at all with the theme of the burglary video, nor the suggestions made in the postevent questions. Subjects were told that as each number–word pair was shown they were to: (1) write down the two-digit number contained on the slide, and (2) write down the first three words that came to their mind when they thought about the word contained on the slide.

Each test took approximately 5 min to complete. Every attempt was made to convince the subjects that these were real experimental tasks by using instructions and answer sheets of comparable complexity and tone as those used in the actual experimental tasks. The tasks were administered in the same order (violence rating, word association) to all subjects.

As in Experiment 1, a 7-min filled delay followed the last subset of postevent questions for all subjects, after which the surprise source memory test was administered.

Source memory test. The test probes and procedure were the same as those used in Experiment 1. The instructions were nearly identical, except for minute changes necessary to facilitate the new answer options (e.g., rather than tell the subjects to circle ‘‘yes’’ in the ‘‘Video’’ column if they remembered the item from the video they were told to ‘‘. . . circle one of the ‘yes’ responses’’).

Subjects were given an answer sheet which contained two columns labeled ‘‘Video’’ and ‘‘Questions.’’ Each column contained 7-option Likert-type scales containing the following responses: ‘‘definitely yes,’’ ‘‘probably yes,’’ ‘‘maybe yes,’’ ‘‘unsure,’’ ‘‘maybe no,’’ ‘‘probably no,’’ ‘‘definitely no.’’ Subjects were told to circle the response in each column that best described the nature of their memory about the source of the information contained in each of the 32 probes.

Results and Discussion

The analyses revealed that the spacing manipulation had no effect on any of the measures of interest (i.e., there were no reliable main effects or interactions involving spacing, all p’s $\geqslant .10$; see the bottom half of Tables 3 and 4, and Table 5). Therefore, we will focus on the effect of the Modality variable in our discussion of the results. For ease of comparison with the results of Experiment 1, we first analyzed the data dichotomized as either ‘‘yes’’ or ‘‘no’’ responses (i.e., summed across definitely, probably, and maybe). Overall, the pattern of results closely replicated those of Experiment 1.

Item Recognition. Once again, as can be seen in the bottom half of Table 3, item recognition was better for repeated items than those

TABLE 5

MEAN PROPORTION OF RECOGNIZED SUGGESTIONS ATTRIBUTED TO THE VIDEO (SOURCE MISATTRIBUTION ERRORS) AS A FUNCTION OF GROUP AND NUMBER OF EXPOSURES IN EXPERIMENT 2

Condition	No. of exposures 1	No. of exposures 3	Repetition effect (3 minus 1)
Single Modality Consecutive	.49	.60	.11
Single Modality Spaced	.51	.61	.10
Mixed Modality Consecutive	.54	.72	.18
Mixed Modality Spaced	.45	.62	.17

suggested once ($F(1, 392) = 154.36$, $MSE = .02$, $p < .0001$). As in Experiment 1, there were no item recognition differences between the Single Modality (*M*'s = .83 and .97 for one and three exposures, respectively) and Mixed Modality subjects (*M*'s = .81 and .95 for one and three exposures, respectively).

Source misattribution errors. Figure 2 depicts conditionalized source misattribution errors collapsed across Spacing Group (although individual group means are reported in Table 5). Inspection of the figure shows that the main results of Experiment 1 were clearly replicated. That is to say, while repeated exposure to suggestion increased errors overall ($F(1,392) = 77.82$, $MSE = .05$, $p < .0001$), the repetition effect was once again greater for Mixed than Single Modality subjects, as evidenced by a reliable Exposure X Modality interaction ($F(1,392) = 5.09$, $MSE = .05$, $p < .05$). Planned comparisons confirmed that there was a significant repetition effect in both Modality Groups ($F(1,197) = 27.34$, $MSE = .04$, $p < .0001$, and $F(1,197) = 52.57$, $MSE = .06$, $p < .0001$, for Single and Mixed Modality, respectively) and that the interaction resulted because there was no difference in error rates for the two groups at the 1-exposure level ($F(1,394) = .002$, $MSE = .16$, $p = 1.0$), but the increase in errors committed by the Mixed Modality subjects at the 3-exposure level approached significance ($F(1,394) = 3.27$, $MSE = .15$, $p = .07$). As in Experiment 1, it is clear that the confusion on the part of the Mixed Modality subjects about whether or not the suggestions appeared in the video was specific to the case in which the suggestions were actually exposed in all three modalities.

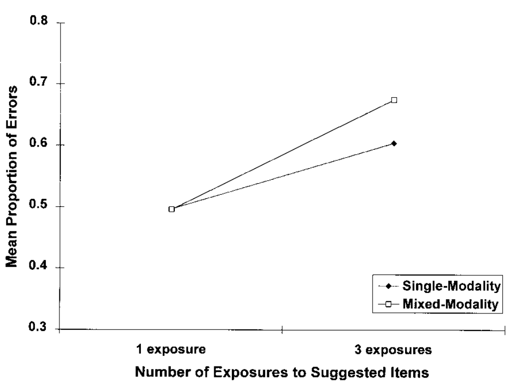

FIG. 2. Mean proportion of recognized suggestions attributed to the video (source misattribution errors) as a function of modality group (i.e., collapsed across Spacing variable) and number of exposures in Experiment 2.

Memory for actual source. Analysis of subjects' memory for actual source (defined as the sum of ''yes'' responses in the ''Questions'' column) revealed once again that repeated exposure increased subjects' ability to remember that the suggestions were in the postevent questions ($F(1,392) = 37.61$, $MSe = .03$, $p < .0001$, see the bottom half of Table 4). However, in contrast to Experiment 1, there was no evidence that contextual variability impaired memory for actual source. The Mixed Modality subjects (*M*'s = .88 and .97 for one and three exposures, respectively) did not differ from the Single Modality subjects (*M*'s = .91 and .96 for one and three exposures, respectively) on accurate source memory, nor did

Modality interact significantly with number of exposures (p's $>$ .05). However, the possibility that ceiling effects may have obscured any differences cannot be ruled out.

Confidence. Although we made no *a priori* predictions concerning the effects of the contextual manipulations on subjects' confidence in their memories for the suggestions, the 7-point confidence scale obviously has the potential to provide data concerning the subjective experience underlying subjects' source decisions. We therefore analyzed source responses in the individual confidence categories separately (i.e., ''maybe yes,'' ''probably yes,'' ''definitely yes''). Analyses of source misattribution errors showed that the Exposures X Modality interaction was not centered in any single confidence level, as it was not mirrored in any of the categories, but rather was the result of the summed ''yes'' responses in the ''Video'' column. Furthermore, there were no significant main effects of Modality in any of the individual confidence categories (all p's $>$ .05). Analyses of memory for actual source by category also revealed no significant differences between the groups (i.e., no interactions or main effects; all p's $<$. 05). Finally, we note that there were no group differences in use of the ''unsure'' option (M's = .16 and .14 at the one and three exposure levels for Single Modality subjects; .18 and .10 for Mixed Modality subjects, p's $\geqslant$ .05).

In summary, the results of Experiment 2 show that one manipulation of contextual variability increased source misattributions and the other did not. Specifically, as in Experiment 1, the contextual variability introduced by changing the modality in which repeated suggestions were encountered increased subjects' tendency to misremember witnessing the suggestions in the video. However, the contextual variability that is presumably caused by separating the repetitions with filled, albeit small (5 min), intervals of time had no effect on subjects' error rate. It may be that this particular manipulation of context variability, being a more indirect method than changing modalities (cf. Bellezza & Young, 1989), was not strong enough to reliably affect subjects' source memory.

General Discussion

In both experiments reported here, repeated exposure to suggestion increased false memory for the suggested events, thus replicating Zaragoza and Mitchell (in press). These results extend those findings, however, by demonstrating that increasing the contextual variability of the repeated exposures can magnify this effect. Subjects who encountered the repeated suggestions in different modalities were more likely to misattribute the suggestions to the video than subjects who received all of them in the same modality. It is surprising that a relatively minor change in context—modality—nevertheless resulted in a reliable increase in errors, especially given that there was already considerable contextual variability across repeated exposures. Recall that in all groups, each exposure to a suggested item occurred in a different question and thereby varied in both linguistic and semantic context. In addition, each exposure was separated by 11 intervening questions. Thus, adding the modality change probably introduced a relatively small increment in contextual variability. That this rather modest manipulation of contextual variability was sufficient to boost errors lends confidence to the conclusion that contextual variability plays a role in the deleterious effects of repeated suggestions. What is not clear from the present study is the extent to which contextual variability might be *required* for the effects of repeated suggestion to be observed. This remains an important question for future research.

Why might increasing the contextual variability of repeated exposures exacerbate subjects' tendency to misremember witnessing the suggested items? In attempting to answer this question, we first consider the potential causes of source misattribution errors in this paradigm. Generally speaking, given that an item is recognized, there are probably two conditions that must jointly be met for a source misattribution to occur: (1) the target memory contains characteristics that are similar to those of a different source and (2) there is no information accessed from memory which prevents subjects from misattributing the item's source. When subjects attempt to answer misleading questions about witnessed events they are likely to think about and visually imagine the events described in the questions. We have proposed that it is this imagined representation of the suggested information that gets confused for an actually perceived event. We have also posited that with repeated exposure to suggestion this image of the suggested event may become increasingly elaborated and detailed (Suengas & Johnson, 1988), thus more closely approximating the amount of sensory and perceptual detail found in memories for witnessed events. Because these imagined representations of the suggested information are not generated intentionally, information about the cognitive operations that went into constructing these images may be lost rather quickly. What is likely to remain, therefore, is simply a memory record of imagined information that is easily confused for information derived from a witnessed event. Note that in this case, accurate memory for having read a suggested item is not very diagnostic with regard to whether or not the item was in the video, since most of the information described in the questions was in the video also.

When the repeated exposures were encountered in different contexts, there were several potential consequences for memory, any or all of which might have served to increase source misattribution errors. One possibility is that subjects' awareness that the suggested items were repeated might have decreased with increases in contextual variability. Recognition that an item is repeated might help to prevent source misattributions, especially in those cases where subjects detect the misinformation as potentially false on the first exposure. If subjects fail to notice that an item is repeated they are less likely to retrieve and later remember the fact that they had earlier questioned the veracity of the misinformation.

Another possibility is that contextual variability led to increases in the suggested items' familiarity for the Mixed Modality Group that were not detected because of ceiling effects in our item recognition measure. If the repeated items evoked an overwhelming sense of familiarity in the Mixed Modality Group, subjects may have taken the high familiarity as an indication that the suggested items must have come from both the video and the questions, regardless of what they actually remembered about the items' source (see also, Whittlesea, 1993; Whittlesea, Jacoby, & Girard, 1990). Such an inference would lead to an increase in errors (see Jacoby, Kelley, & Dywan, 1989, for a review of misattributions of familiarity).

A third possibility is that contextual variability may have resulted in the ''generification'' (cf. Watkins & Kerkar, 1985) of subjects' memory for the suggested items, whereby subjects could no longer remember specific information about their source. Although the near ceiling levels of actual source memory performance in the Mixed Modality Groups may seem to imply the contrary—that subjects had excellent memory for the suggested items' true source—this is not necessarily the case. To correctly identify the suggestions as having occurred in the postevent questions the subjects merely had to determine that they had occurred recently, in the latter half of the experiment. This should have been easy given that the questioning phase occurred shortly before the test. Moreover, the fact that the repeatedly suggested items were probably highly familiar should have biased subjects to assume that they had been presented quite recently (see Whittlesea, 1993, for evidence consistent with this prediction). Thus, it is possible that contextual variability led to much less differentiated information about the suggested items' source, but that this was not detected by our source measure. The absence of discrete source information in combination with the suggested items' high familiarity and misleading sensory/perceptual characteristics may have led to an increase in source misattribution errors.

Determining the exact mechanism(s) by which the effects of contextual variability come about cannot be determined without further research. What is lacking is a precise characterization of the nature of subjects' memory for the source of the suggested items. Perhaps more fine-grained measures of subjects' recollective experience can accomplish this goal. Optimally, such a measure would identify the aspects of subjects' phenomenal experience that lead them to make these errors.

In addition to their implications for false memory, the present results also provide new information about repetition and contextual variability. Prior research on contextual variability in memory for repeated information has emphasized its beneficial effects on item recall and recognition (see Greene, 1992, for a review). The results of the present study show, however, that if the learner's goal is to remember the precise origin of this information, such variability may prove harmful. Specifically, in these experiments, increasing the contextual variability between exposures to suggested items impaired subjects' ability to accurately discriminate the precise source of these items—the postevent questions—as evidenced by their inability to avoid misattributing these items to the video also. This divergence in the effects of contextual variability underscores the notion that source memory and item memory can invoke different processes and rely on different sorts of information (Johnson et al., 1993).

Finally, we note the potential practical ramifications of this research. Awareness of the important social and legal implications of research on false memory has recently led investigators to emphasize a high degree of ecological validity in designing their studies. Clearly, such a research strategy can provide unique information bearing on these real-world concerns. However, this approach needs to be complemented by more analytic studies which seek to understand the processes by which false memories are constructed. To date, relatively little attention has been given to understanding the individual factors that contribute to the formation of these memories (but see the work of Hyman and colleagues for an exception, e.g., Hyman & Pentland, this issue). We have recently identified one such factor, namely, repeated exposure to suggestion (Zaragoza & Mitchell, in press). The present study advances our understanding of how repetition affects the creation of false memories by identifying contextual variability as a mediator in its effects. The success of this initial endeavor gives us confidence that systematic attempts to investigate the mechanisms by which false memories arise will ultimately lead to a more complete understanding of this phenomenon.

APPENDIX A: MISLEADING POSTEVENT SUGGESTIONS

thief wore gloves

thief pulled down a *window shade*

thief stole a *ring*

driver smoked a *cigarette*

neighbor's name was *Mrs. Anderson*

barking dog

thief had a *gun*

thief put his *seatbelt* on

police officer had a *Coke*

police officer said driver was *DWI*

car *jumped a curb*

police said they'd *shoot*

REFERENCES

ACKIL, J. K., & ZARAGOZA, M. S. (1995). Developmental differences in suggestibility and memory for source. *Journal of Experimental Child Psychology,* **60,** 57–83.

ARKES, H. R., HACKETT, C., & BOEHM, L. (1989). The generality of the relation between familiarity and judged validity. *Journal of Behavioral Decision Making,* **2,** 81–94.

BACON, F. T. (1979). Credibility of repeated statements: Memory for trivia. *Journal of Experimental Psychology: Human Memory and Learning,* **5,** 241–252.

BEGG, I. M., ANAS, A., & FARINACCI, S. (1992). Dissociation of processes in belief: Source recollection, statement familiarity, and the illusion of truth. *Journal of Experimental Psychology: General,* **121,** 446–458.

BEGG, I., & ARMOUR. (1991). Repetition and the ring of truth: Biasing comments. *Canadian Journal of Behavioural Science,* **23,** 195–213.

BELLEZZA, F. S., & YOUNG, D. R. (1989). Chunking of repeated events in memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition,* **15,** 990–997.

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 and Cognition,* **21,** 40–54.

CECI, S. J., LOFTUS, E. F., LEICHTMAN, M. D., & BRUCK, M. (1994). The possible role of source misattributions in the creation of false beliefs among preschoolers. *The International Journal of Clinical and Experimental Hypnosis,* **42,** 304–320.

CECI, S. J., CROTTEAU HUFFMAN, M. L., SMITH, E., & LOFTUS, E. F. (1994). Repeatedly thinking about non-events. *Consciousness and Cognition,* **3,** 388–407.

GREENE, R. L. (1992). *Human memory: Paradigms and paradoxes.* Hillsdale, NJ: Erlbaum.

HASHER, L., GOLDSTEIN, D., & TOPPINO, T. (1977). Frequency and the conference of referential validity. *Journal of Verbal Learning and Verbal Behavior,* **16,** 107–112.

HYMAN, I. E., JR., & BILLINGS, F. J. (1995). *Individual differences and the creation of false childhood memories.* Unpublished manuscript.

HYMAN, I. E., JR., HUSBAND, T. H., & BILLINGS, F. J. (1995). False memories of childhood experiences. *Applied Cognitive Psychology,* **9,** 181–197.

HYMAN, I. E., JR., & PENTLAND, J. (1996). The role of mental imagery in the creation of false childhood memories. *Journal of Memory and Language,* **35,** 101–117.

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

JAMES, W. (1890/1918). *The principles of psychology (Vol. 1).* New York: Dover Publications.

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

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

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

LINDSAY, D. S., & READ, J. D. (1994). Psychotherapy and memories of childhood sexual abuse: A cognitive perspective. *Applied Cognitive Psychology,* **8,** 281–338.

LOFTUS, E., & KETCHAM, K. (1994). *The myth of repressed memory.* New York: St. Martin's Press.

MACLEOD, C. M. (1988). Forgotten but not gone: Savings for pictures and words in long-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition,* **14,** 195–212.

RICCIO, D. C., ACKIL, J., & BURCH-VERNON, A. (1992). Forgetting of stimulus attributes: Methodological implications for assessing associative phenomenon. *Psychological Bulletin,* **112,** 433–445.

ROVEE-COLLIER, C. (1991). The ''memory system'' of prelinguistic infants. *Annals of the New York Academy of Sciences,* **608,** 517–542.

SCHWARTZ, M. (1982). Repetition and the rated truth value of statements. *American Journal of Psychology,* **95,** 393–407.

SUENGAS, A. G., & JOHNSON, M. K. (1988). Qualitative effects of rehearsal on memories for perceived and imagined complex events. *Journal of Experimental Psychology: General,* **117,** 377–389.

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

WARREN, A. R., & LANE, P. (1995). Effects of timing and type of questioning on eyewitness accuracy and suggestibility. In M. S. Zaragoza, J. R. Graham, G. C. N. Hall, R. Hirschman, & Y. S. Ben-Porath (Eds.), *Memory and testimony in the child witness* (pp. 44–60). Thousand Oaks, CA: Sage.

WATKINS, M. J., & KERKAR, S. P. (1985). Recall of a twice-presented item without recall of either presentation: Generic memory for events. *Journal of Memory and Language,* **24,** 666–678.

WHITTLESEA, B. W. A. (1993). Illusions of familiarity. *Journal of Experimental Psychology: Learning, Memory, and Cognition,* **19,** 1235–1253.

WHITTLESEA, B. W. A., JACOBY, L. L., & GIRARD, K. (1990). Illusions of immediate memory: Evidence of an attributional basis for feelings of familiarity and perceptual quality. *Journal of Memory and Language,* **29,** 716–732.

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

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

(Received July 21, 1995)
(Revision received December 13, 1995)