Journal of Experimental Psychology: Applied
2002, Vol. 8, No. 2, 127–142

1076-898X/02/$5.00 DOI: 10.1037//1076-898X.8.2.127

New Evidence on the Suggestibility of Memory: The Role of Retrieval-Induced Forgetting in Misinformation Effects

Jo Saunders and Malcolm D. MacLeod
University of St. Andrews

Extending recent work that has demonstrated that the act of remembering can result in the inhibition of related items in memory, the present research examined whether retrieval-induced forgetting could provide a mechanism for explaining misinformation effects. Specifically, the authors found in their first study that the inhibition of critical items rendered the recollection of postevent information more likely in a subsequent test of memory. The authors established in their second study that when guided retrieval practice and final recall tests were separated by 24 hr, retrieval-induced forgetting failed to emerge and misinformation effects were absent. In contrast, a delay of 24 hr between initial encoding and guided retrieval practice produced not only retrieval-induced forgetting but also misinformation effects.

> In the practical use of our intellect, forgetting is as important a function as remembering.
>
> —James, 1905, p. 300

One of the most intriguing debates in psychology concerns the permanence of memory. Since the days of William James, psychologists have theorized on its properties and, in particular, whether memories that are currently inaccessible are truly lost for all time. In the quest to shed light on this complex issue, psychologists have left few mental stones unturned. Yet, despite this prodigious level of research activity, our understanding of the interplay between old and new memories remains relatively limited. The lack of specificity regarding the mechanisms by which updating is accomplished and the form of these updated memorial representations, in turn, continue to place limits not only on the theoretical modeling of memory but also the likely resolution of everyday memory problems. Among those issues that stand to benefit from an enhanced understanding of how memory updating is achieved include the assessment of eyewitness reliability, the role of memory in adjustment to trauma, and the accuracy of recovered memories for childhood sexual abuse.

One area of psychological research that has traditionally been concerned with both applied and theoretical aspects of memory has been the study of eyewitness suggestibility. Elizabeth F. Loftus and others have focused much of their research efforts on how eyewitnesses report honest errors following the introduction of misleading postevent information. This pioneering work led Loftus to advance the thesis that the introduction of inconsistent information presented subsequent to original encoding can overwrite memory for original material (Loftus, 1979; Loftus & Loftus, 1980; Loftus, Miller, & Burns, 1978). This form of destructive updating of memory, in turn, is thought to be responsible for what is termed the *misinformation effect* (i.e., the tendency for people to recall misleading postevent information in preference to original material).

Loftus and others have argued (Loftus, Schooler, & Wagenaar, 1985; Metcalfe, 1990) that disruption or damage to the original memory trace consequent on the introduction of inconsistent information can result in a blended representation of the original and postevent material. Depending on the nature of the conditions used at test, blended representations would tend to favor different outcomes. Under test conditions where the choice is between the original material and misinformation, misinformation is likely to prevail (e.g., Loftus et al., 1978). In contrast, where the choice is between previously unseen material and original material, the original would tend to be favored. McCloskey and Zaragoza (1985) argued that because memory performance under such conditions is little different from controls who have not been misled, there is no conclusive evidence that misinformation alters memory for original material. Loftus et al. (1985), in contrast, argued that the procedure adopted by McCloskey and Zaragoza is insensitive to memory blending. In other words, the original material would still be more similar to the blended representation than any previously unseen material (see also, Belli, 1989). Finally, under conditions where one of the options represents a compromise between the original material and misinformation (e.g., the option of bluish green where the original item was green and the misinformation was blue), the compromise option would tend to be chosen (e.g., Loftus, 1977).

Composite recollections have considerable intuitive explanatory power. Not only can they account for many of the findings that exist in the misinformation literature, but they can also explain some of the common errors that arise in autobiographical memory (Neisser, 1981). Many of us, for example, have experienced recalling much-enjoyed activities on a childhood holiday only to find that the events eagerly brought to mind had actually comprised a number of such holidays. Clearly, it is undeniable that such com-

Jo Saunders and Malcolm D. MacLeod, School of Psychology, University of St. Andrews, Fife, Scotland.

We acknowledge the support of the Economic and Social Research Council, who awarded a studentship to Jo Saunders under the supervision of Malcolm D. MacLeod. We also thank Mike Anderson for thoughtful comments on an earlier version of this article.

Correspondence concerning this article should be addressed to Malcolm D. MacLeod, Social and Applied Cognition Lab, School of Psychology, University of St. Andrews, Fife, Scotland KY16 9JU. E-mail: mdm@st-andrews.ac.uk

posite recollective experiences exist but their presence does not necessarily imply that blending or superimposed representations are responsible (see Schooler & Tanaka, 1991). Indeed, the actual empirical evidence for memory blending is not particularly strong. Using a range of options presented via a color wheel, for example, Belli (1988) demonstrated that where participants were misled, the choice of color suggested by the misinformation was chosen in preference to any compromise color suggested by the combination of misinformation and original material.

Other researchers have adopted a different approach to explaining misinformation effects. Rather than arguing that the introduction of inconsistent postevent information alters original memory, they argue that both memories coexist and that the conditions that prevail at retrieval determine the memory to be accessed. In other words, misinformation effects can be explained in terms of retrieval accessibility without recourse to blending or the superimposition of memories (Bekerian & Bowers, 1983; Morton, Hammersley, & Bekerian, 1985). The basic idea here is that memory for the original material survives intact but is less accessible during retrieval because of interference set up by memory for the more recent misinformation (see also, Christiaansen & Ochalek, 1983). Other coexistence explanations for the misinformation effect focus on source confusion and report criteria. Lindsay and Johnson (1989b), for example, demonstrated that misinformation effects can be elicited even under conditions in which participants are presented with a narrative description of a scene containing misleading suggestions before viewing the scene to which the narrative referred. They argued that this "reversed suggestibility effect" is difficult to reconcile with a destructive-updating explanation that is dependent on newer memories updating older ones, and argued that misinformation effects may be due to confusion over memory source. When people are encouraged to attend to information concerning the source of memories, misinformation effects can be eliminated (Lindsay & Johnson, 1989a). Indeed, it has been shown that such source misattribution errors not only give rise to impaired memory (Lindsay, 1990), but that people can mistakenly believe that they had seen the erroneous information during original encoding (Weingardt, Loftus, & Lindsay, 1995).

Notwithstanding these unresolved theoretical debates, the phenomenon of misinformation effects has had an unprecedented impact on lay perceptions concerning the permanence and reliability of memory. Yet, if we are to understand the consequences of misinformation effects and how best to prevent their occurrence, it is important that we refocus our attention on the underlying mechanisms. As indicated above, this is not straightforward. Under some conditions, the most expeditious explanation for the misinformation effect would seem to involve some form of destructive updating of original memory (Loftus, 1979; Loftus et al., 1978). Under other circumstances, however, misinformation effects can best be understood in terms of retrieval competition (Bekerian & Bowers, 1983; Morton et al., 1985) or the misattribution of source memories (Lindsay & Johnson, 1989a). Some studies indicate the coexistence of memories for both original and inconsistent material (Bekerian & Bowers, 1983; McCloskey & Zaragoza, 1985), whereas others indicate that any conflict between memories is resolved at the time of comprehending the postevent information rather than responding at test with two conflicting representations (Loftus, Donders, Hoffman, & Schooler, 1989).

This disparate set of findings suggests two possibilities. The first is that there may be many routes by which misinformation effects are produced (see Belli, 1989)—a possibility also acknowledged by Loftus and colleagues (e.g., Loftus & Hoffman, 1989; Loftus, Korf, & Schooler, 1989). Indeed, under some circumstances, we know that factors such as prior expectations (Pezdek, 2001) and demand characteristics (Weinberg, Wadsworth, & Baron, 1983) can also influence the report of misinformation. However, a second possibility is that the findings in the misinformation literature fail to provide a single coherent picture because we have yet to identify the underlying mechanism responsible for many of these effects. This latter option is particularly attractive given that, if we can identify the underlying mechanism, there opens up the possibility of generating greater predictive power than that currently permitted by blending, response competition, or source misattribution explanations. If this latter option is the case, what kind of mechanism could account for the coexistence of memories for original and postevent information under certain circumstances, yet the inaccessibility of original memory under other conditions? The present article addresses this issue.

We argue that one of the potential mechanisms underlying misinformation effects is *retrieval-induced forgetting* (i.e., the forgetting of information as a function of remembering other related material; see M. C. Anderson, Bjork, & Bjork, 1994; M. C. Anderson & Spellman, 1995). This effect is thought to be due to the active suppression or inhibition of unwanted material at time of retrieval. Typically, the retrieval cues we use are insufficiently specified and therefore, not only do we access the desired material but also related unwanted material. By inhibiting such related items in memory, we can effect a reduction in retrieval competition and promote the retrieval of desired information. The most compelling evidence that inhibition is responsible for this phenomenon comes from studies demonstrating cue-independent forgetting (M. C. Anderson & Green, 2001; M. C. Anderson & Spellman, 1995). The rationale here is that if inhibition is responsible then recall for inhibited items should be worse than recall for noninhibited items, irrespective of whether the items are tested with the original cue or whether novel cues are used. If mechanisms such as associative interference were responsible then we would expect recall to be enhanced where novel cues are used. The fact that recall performance for such items remains below baseline, irrespective of the cues used during retrieval, eliminates associative interference as a possible explanation and provides strong support for the existence of an inhibitory control mechanism (M. C. Anderson & Green, 2001; M. C. Anderson & Spellman, 1995).

Retrieval-induced forgetting has a number of important associated properties. First, unlike directed forgetting effects (R. A. Bjork, 1989; Johnson, 1994), explicit cues to forget are not required. Rather, such forgetting is considered to be the undirected response to dealing with unwanted related information competing for retrieval (M. C. Anderson, Bjork, & Bjork, 1994, 2000; M. C. Anderson & McCulloch, 1999; M. C. Anderson & Spellman, 1995). Second, recent work in this field has suggested that this effect is relatively easy to elicit (Macrae & MacLeod, 1999)—an important feature if it is to have relevance beyond the confines of the laboratory—and that, for certain kinds of information, it is also nonpermanent (MacLeod & Macrae, 2001). This latter feature is of particular importance given that many of the goals in everyday life are in a constant state of flux. Thus, for example, the telephone

number that is redundant today may well prove to be of importance at a later date. To render certain kinds of information permanently unavailable to conscious inspection would only be adaptive if one's goals never changed or were always novel. Third, these adaptive properties can have undesirable consequences, at least under certain task conditions. Macrae and MacLeod (1999) demonstrated some of the consequences of this adaptive mental process in the realm of social cognition (see also, E. L. Bjork, Bjork, & MacLeod, 2002, for a discussion). Fourth, the potential for retrieval-induced forgetting has already been demonstrated in studies of eyewitness memory (MacLeod, 2002; Shaw, Bjork, & Handal, 1995). The present article extends this body of work by exploring the extent to which retrieval-induced forgetting contributes to the production of misinformation effects.

If we consider the standard misinformation paradigm (e.g., Loftus et al., 1978), participants are given a postevent questionnaire that typically asks about the previously witnessed event. Such questionnaires, however, do not constitute an exhaustive retrieval of the information known about the event but rather focus on a subset of items. We also know that it is the act of retrieval that sets up the inhibition of related unpracticed items (M. C. Anderson et al., 2000); that is, the mere presentation of information in the absence of retrieval practice is not sufficient to produce retrieval-induced forgetting. The retrieval of a subset of items about the event, therefore, is likely to produce inhibition of other related pieces of information about the event that have not been the subject of the questionnaire. Where misinformation is then introduced about one of these inhibited items, a subsequent test of memory for the critical item should result in a preference for the postevent misinformation, as the original is no longer available to conscious inspection. For example, let us consider the following critical question used by Loftus et al. (1978): "Did another car pass the red Datsun while it was stopped at the stop sign?" Here there is a request to retrieve information about a second possible car. It does not require the retrieval of information about the type of sign at the crossroads. Assuming that no other question focused on the type of sign, when participants' memories are subsequently tested, memory that it was a "Yield" sign would be inhibited as part of the original memory for the event, leaving only the postevent memory available for retrieval (i.e., "Stop" sign). In other words, the postevent questionnaire used in the standard misinformation study in order to introduce misinformation could set up the conditions necessary for the forgetting of the critical unpracticed detail (i.e., the production of retrieval-induced forgetting) which, in turn, facilitates the report of misinformation about the critical item.

It is important to note that we are not suggesting that the misinformation effect reported by Loftus and others (e.g., Loftus et al., 1978) is an artifact of the paradigm used to investigate it; indeed, quite the converse. Rather, we are attempting to identify the conditions under which misinformation effects are likely to be produced. The fact that retrieval-induced forgetting is likely to occur where questions have been asked about a subset of items known about an event is of theoretical interest and potentially of considerable practical importance in understanding when misinformation effects are most likely to occur. For example, most police interviews do not involve exhaustive retrievals for information about a witnessed event for a variety of reasons (MacLeod, 2002; Shaw et al., 1995). Sometimes this may be because the investigating police officer is principally concerned with current investigative objectives that, in turn, determine those aspects of the event that are considered most important. In other instances, there may simply be insufficient time or resources to carry out exhaustive interviews, or the police officer may fail to grasp the complexity of what has occurred. Shaw et al. (1995) also pointed out that it is conceivable that those details that were not the subject of initial retrieval practice could in time become critical aspects of a case and that these details, in turn, could be more poorly recalled during any subsequent retrieval of information (e.g., during trial proceedings). We would add that the incomplete retrieval of information about an event could produce those conditions that promote the likelihood of misinformation inadvertently being reported about items that have not been the subject of initial retrieval practice. Therefore, the fact that the questionnaire used to introduce the misinformation in the standard misinformation paradigm parallels the incomplete retrieval of information in police interviews simply strengthens the relevance of misinformation studies to our understanding of real-life issues.

In retrieval-induced forgetting studies, four phases are typically used comprising study, retrieval practice, distractor task, and final test (see M. C. Anderson et al., 1994). In the study phase, participants are presented with a series of category-exemplar pairs and instructed to memorize the items (e.g., *drink–beer, fish–salmon*). Directed retrieval practice is then performed on half the studied exemplars from half the categories by completing a series of cued-stem tests (e.g., *drink–be__*). Each of these cued exemplars is typically presented on three occasions. Following a distractor task, participants are then given a recall test in which they are cued with each category name (e.g., *drink, fish*) and asked to recall all the exemplars originally presented to them. In such studies, the recall performance of three types of exemplars is assessed: Rp+ items (i.e., practiced exemplars from practiced categories, *beer*); Rp− items (i.e., unpracticed exemplars from practiced categories, *sherry*); Nrp items (i.e., unpracticed exemplars from unpracticed categories, *halibut*). If remembering some items inhibits the memorability of related material, then participants should show poorer recall performance for unpracticed items from the same category as the practiced items than for unpracticed items from the previously unpracticed category (i.e., they should have greater difficulty recalling Rp− items than Nrp items).

Using a variation of this basic paradigm (M. C. Anderson et al., 1994), the present inquiry aims to establish whether retrieval-induced forgetting can account for the production of misinformation effects. Typically, studies that have explored retrieval-induced forgetting effects have reported differences between Rp− and Nrp recall performance in the region of −.10 to −.19 (see M. C. Anderson et al., 1994; M. C. Anderson & Spellman, 1995; MacLeod, 2002; MacLeod & Macrae, 2001; Macrae & MacLeod, 1999). If our contention is correct, we can make a number of specific predictions. Specifically, we would expect to find misinformation effects where misleading postevent information has been introduced about Rp− items (i.e., items inhibited as a result of retrieval practice on other related items), but absent where introduced about noninhibited items (i.e., Rp+ or Nrp items). We could also expect that, as recent research (MacLeod & Macrae, 2001) has indicated that such inhibition is transient (at least for some kinds of material), misinformation effects would only be evident where the critical items remain inhibited. MacLeod and Macrae (2001) demonstrated that inhibition is still active where

there is a 24-hr delay between study and retrieval practice but not where there is a similar delay between retrieval practice and final test. Thus, following the introduction of misleading information, we could expect to find misinformation effects occurring only in the former of these two conditions, that is, where inhibition is still active. We tested these predictions in the two studies reported herein.

Study 1: Retrieval-Induced Forgetting and Misinformation Effects

Method

Participants and Design

One hundred undergraduate students (42 men and 58 women) participated on a voluntary basis in this study. The experiment had a single factor (misinformation item [Mis]: MisRp+, MisRp−, MisNrp, or control) between-subjects design where postevent misinformation was introduced about either an Rp+, Rp−, or Nrp item, respectively. A control condition was also included where no relevant retrieval practice had taken place. Each condition included 25 participants.

Procedure and Stimulus Materials

Participants arrived at the laboratory individually or in groups of up to four, were greeted by a female experimenter, and randomly assigned to one of the testing conditions. The basic procedure followed Shaw et al. (1995) and MacLeod (2002), but with the addition of phases that (a) permitted the introduction of misleading postevent information and (b) the assessment of the misinformation effect. Participants were informed that they were to take part in a memory task and were instructed to read two narratives about two separate burglaries. The order of presentation of the two narratives had been counterbalanced throughout. Information about the burglaries was contained within an experimental booklet that also contained a number of distractor tasks in addition to appropriate retrieval-practice questions. Participants were prompted through each phase of the booklet by the experimenter (see Figure 1 for a summary of the experimental procedure). The first part of each narrative contained scene-setting information about when and where the crime had occurred. One narrative concerned the theft of items from the Jones' [sic] house and described how the Jones' daughter had arrived home to find the house had been broken into. Ten items were described as having been stolen from the house (i.e., Game Boy, sunglasses, mobile phone, painting, binoculars, wristwatch, printer, television, coffee maker, china plate). Each item was presented embedded within a set of sentences describing where the stolen item had originally been located within the house (e.g., "The television had been in the sitting room, which is at the front of the house. It was sitting in the corner of the room. The remote control for it hadn't been taken."). All the items chosen for each household were considered believable as potential stolen items in a burglary. No attempt was made to match any of the items in terms of semantic similarity across households, although some matching may have arisen simply as a function of creating two sets of stealable objects (see *Additional Analyses* section for further discussion).

Participants were informed that the underlined words represented the stolen items. Each item was presented on a separate page of the booklet. Participants were instructed to turn over to the next page only on hearing an audible beep emitted by an electronic metronome (5-s intervals). On completing the first narrative, participants were instructed to read the next narrative about a burglary at the Smith's [sic] house. This similarly detailed a number of items that had been stolen (i.e., computer, video recorder, telescope, rollerblades, necklace, Discman, camcorder, leather coat, crystal vase, microwave) and presented in the same fashion (see Appendix A for full details of the narratives used). The stolen items were presented in blocked format (i.e., all items about the Jones' house followed by the Smith's house or vice versa) and their presentation fully randomized within each block. The information sets for each house were divided into two subgroups (each containing five items) for the purpose of creating a practiced (i.e., Rp+) and an unpracticed (i.e., Rp−) set of items for each theft (see M. C. Anderson et al., 1994). The contents of each subset were randomized for each participant.

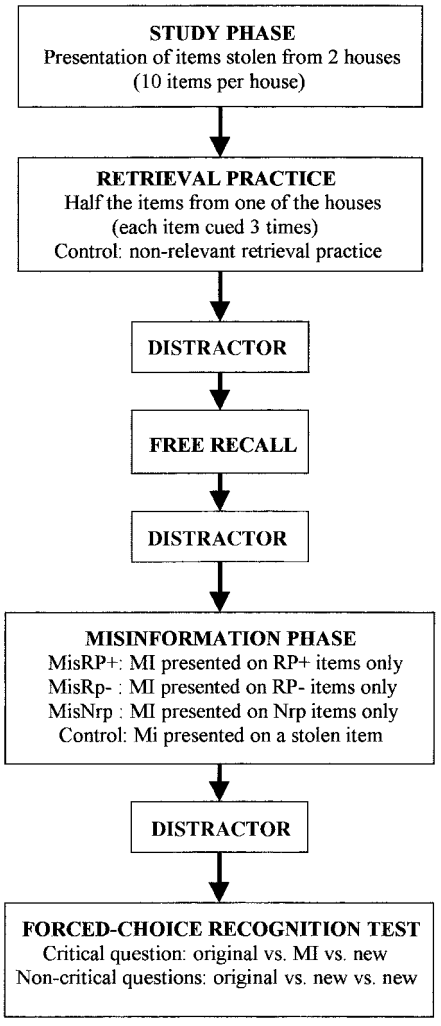

Figure 1. Procedure for Experiment 1.

Immediately after the study phase, participants in the experimental conditions were presented with a series of questions about one of the subsets of stolen items from one of the houses, thereby creating Rp+, Rp−, and Nrp item sets. This series of questions comprised three sets of questions about five of the stolen items. Thus, each participant received a total of 15 retrieval cues. Each question set increased in difficulty as participants progressed through the booklet following the procedure outlined in Shaw et al. (1995) and based on previous research that demonstrated this procedure maximized practice effects (Landauer & Bjork, 1978). Question-set difficulty had been determined in earlier pilot work, and counterbalancing ensured that each of the items appeared equally often as Rp+, Rp−, and Nrp items. Participants in the control condition did not receive retrieval practice about any of the stolen items. Rather, they received retrieval practice for the names of capital cities (e.g., the capital city of Cuba is Ha_____; the capital city of Bulgaria is So_____, etc).

Following the retrieval practice phase, participants were required to perform a distractor task that involved participants writing down the names of 10 animals for each letter of the alphabet. Participants were given 5 min to accomplish this task. No participant successfully completed it within the time set. Following this, participants were instructed to recall all the items they could remember about both burglaries. This free-recall task served as a manipulation check that retrieval-induced forgetting had actually occurred. Following recall, participants performed a further distractor task for 2 min (i.e., "write down as many objects as possible that are black, wooden, blue, and round for each letter of the alphabet"). Again, no participant completed the task within the time allotted. Subsequent to this, participants were presented with 12 additional questions about one of the burglaries (see Appendix B). Participants in the MisRp+ and MisRp− conditions received questions concerning the practiced household only, whereas participants in the MisNrp condition received questions only about the nonpracticed household. One of the questions presented contained a piece of misinformation about one of the stolen items presented in the original study phase. (e.g., *necklace* was replaced with *earrings*). Thus, depending on treatment condition, participants received an erroneous piece of information about an Rp+, or Rp−, or Nrp item. The misinformation item chosen was semantically related in each case to the critical item. All the questions used in this phase of the study referred to details that were neither the subject of the retrieval practice phase or formed the basis of the final memory test. Participants in the control condition also received a piece of misinformation about one of the thefts, thereby providing a baseline misinformation effect in the absence of relevant retrieval practice. Only one misinformation item was incorporated into each set of 12 questions so as not to arouse suspicions about the nature of the study. To minimize possible item effects, we selected four items in each household as possible critical items against which semantically related misinformation would be introduced. The critical items chosen were counterbalanced throughout for each condition.

Participants were then presented with a further distractor task in which participants were required to write down the names of 10 countries for each letter of the alphabet. Five minutes were allotted for this task, and no participant completed the task in the time allowed. The final test comprised a series of multiple-choice questions in order to test memory for the stolen items. Possible answers to each question comprised the originally presented item plus two erroneous items. The critical question contained the original material, the erroneous misinformation plus one new wrong item. This critical question was used to determine whether a misinformation effect was present or not. On completion of the recall test, participants were debriefed, thanked for their participation, and dismissed.

Results and Discussion

The retrieval practice success rates for each of the experimental conditions were as follows: 90.6%, 88.8%, and 87.7% for the MisRp+, MisRp−, and MisNrp conditions, respectively. A set of manipulation checks was undertaken to ensure that retrieval practice had produced the anticipated effects. These effects are important to establish if we are to provide an adequate test of our hypotheses that link the inhibition of information with an increased likelihood of recalling misinformation for inhibited versus noninhibited items. Data displayed in Table 1 reveal that, across the various conditions, recall performance was .43 for unpracticed items from the unpracticed set but only .26 for unpracticed items from the practiced set. This pattern suggests that retrieval-induced forgetting had occurred. Indeed, the magnitude of the difference in recall performance between Rp− and Nrp items (−.17) was comparable to that reported in other studies of retrieval-induced forgetting (M. C. Anderson et al., 1994; M. C. Anderson & Spellman, 1995; MacLeod, 2002; MacLeod & Macrae, 2001; Macrae & MacLeod, 1999).

Table 1
Mean Recall Performance as a Function of Item Type—Study 1

	Item type			Difference score	
Condition	Rp+	Rp−	Nrp	Rp+ − Nrp	Rp− − Nrp
Rp+ misinformation	.85 (.12)	.25 (.19)	.40 (.23)	.45	−.15
Rp− misinformation	.82 (.17)	.27 (.13)	.44 (.18)	.55	−.17
Nrp misinformation	.83 (.17)	.26 (.15)	.46 (.16)	.57	−.20
Mean	.83 (.15)	.26 (.15)	.43 (.19)	.57	−.17

Note. Standard deviations are enclosed in parentheses. Rp+ items were practiced items from the practiced set, Rp− items were unpracticed items from the practiced set, and Nrp items were unpracticed items from the unpracticed set. Control condition: $M = .56$, $SD = .12$.

A single factor (item type: Rp+, or Rp−, or Nrp) within-subjects analysis of variance (ANOVA) revealed a significant main effect of item type for the MisRp+ condition, $F(2, 48) = 79.13$, $p < .001$, $MSE = 0.039$. Cohen's f was calculated as an unbiased estimate of effect size (Cohen, 1988). This indicated the presence of a large effect for this condition (Cohen's $f = 1.81$). Using Holm's sequential Bonferroni approach to control for familywise error rate, a series of post hoc paired samples t tests indicated the presence of both facilitatory (i.e., Rp+ > Nrp), $t(24) = 8.34$, $p < .001$, and retrieval-induced forgetting effects (i.e., Rp− < Nrp), $t(24) = -2.70$, $p < .05$. Participants also recalled significantly more Rp+ items than Rp− items, $t(24) = 15.67$, $p < .001$.

Similar patterns were observed in the data for the remaining experimental conditions. Specifically, single factor (item type: Rp+, or Rp−, or Nrp) within-subjects ANOVAs revealed significant effects of item type for both the MisRp− and MisNrp conditions, $F(2, 48) = 85.01$, $p < .001$, $MSE = 0.024$, Cohen's $f = 1.88$; and, $F(2, 48) = 67.10$, $p < .001$, $MSE = 0.031$, Cohen's $f = 1.68$, respectively. Again, the effect sizes were large for both conditions (Cohen, 1988, p. 284–285). Using Holm's sequential Bonferroni approach, post hoc paired samples t tests revealed that participants recalled significantly more Rp+ items than Nrp items for both the MisRp− and MisNrp conditions, $t(24) = 7.65$, $p < .001$; and, $t(24) = 6.47$, $p < .001$, respectively (i.e., the presence of facilitatory effects). Similarly, recall performance for Rp− items was found to be significantly lower than that for Nrp items in both the MisRp− and MisNrp conditions, $t(24) = -4.56$, $p < .001$; and, $t(24) = -4.04$, $p < .001$, respectively (i.e., the presence of retrieval-induced forgetting effects). Participants also recalled significantly more Rp+ items than Rp− items in both MisRp− and MisNrp conditions, $t(24) = 13.12$, $p < .001$; and, $t(24) = 13.30$, $p < .001$, respectively.

These manipulation checks serve to establish that, for all three experimental conditions, there are clear indications that the retrieval practice procedure has resulted in significantly poorer recall performance for the unpracticed items from the practiced set compared with the unpracticed items from the unpracticed set (i.e.,

Rp− < Nrp), as anticipated in our hypotheses. Additionally, a series of independent *t* tests confirmed that Nrp recall performance in each of the experimental conditions (overall $M = 0.43$; see Table 1) was significantly lower than that achieved in the control condition ($M = 0.56$, $p < .001$). Thus, the observed difference between Rp− and Nrp recall performance in all three experimental conditions is due to a real drop in the recall performance for Rp− items rather than enhanced recall for Nrp items.

Having demonstrated the presence of retrieval-induced forgetting in the three experimental conditions, we now turn to consider the effects of introducing misinformation about items that had been inhibited as a result of the retrieval practice procedure. The principal comparison of interest here is between the level of misinformation reported by participants in the MisRp− condition (i.e., where misinformation had been introduced about an inhibited item) and that reported by participants where misinformation had been introduced about noninhibited items (i.e., in the MisRp+, MisNrp, and control conditions). Table 2 indicates that when misinformation is introduced, it has a larger effect on unpracticed items from the practiced set (MisRp− condition: 60% of participants chose the misinformation) than for unpracticed items from the unpracticed set (MisNrp condition: 20% of participants chose the misinformation). Thus, it would appear that inhibiting an item about which misleading information is subsequently introduced greatly facilitates the likelihood of reporting the misinformation during final test. Participants in the MisRp− condition were two to four times more likely to make an error on a critical item than were participants in other conditions, and almost six times more likely to make an error on critical than on noncritical items. In addition, participants in the MisRp− condition who did not show a retrieval-induced forgetting effect (i.e., a significant difference between Rp− and Nrp recall performance) had a likelihood of only .17 ($n = 6$) of reporting the misinformation item—a level comparable to that observed in the noninhibited item conditions. Participants in the MisRp− condition who exhibited retrieval-induced forgetting, in contrast, had a likelihood of .74 ($n = 19$) of reporting the misinformation on the critical item.

Chi-square analysis confirmed that there was no significant difference in the proportion of participants reporting misinformation in either the MisRp+ or MisNrp conditions, compared with controls where no relevant retrieval practice had taken place (see Table 2). In contrast, participants in the MisRp− condition reported significantly more misinformation than did participants in the control condition, $\chi^2(1, N = 50) = 5.25$, $p < .05$. We computed effect size using phi ($\Phi = .32$), which indicated the presence of a medium-sized effect. It is also worth noting that, overall, there was no significant difference in the final test performance across conditions for noncritical items (85%, 89%, 86%, and 91% correct for the MisRp+, MisRp−, MisNrp, and control conditions, respectively). Thus, it would appear unlikely that the misinformation effect produced in the MisRp− condition was due to poorer overall recall performance at time of test.

Table 2
Likelihood of Participants Choosing Correct, Misinformation, and New Erroneous Information During Forced-Choice Recognition Test—Study 1

Condition	Critical item		Noncritical item	95% confidence interval
	Misinfo.	New error	Error	
MisRp+	.16	0	.15	.03–.37
MisRp−	.60	0	.11	.39–.81
MisNrp	.20	0	.14	.03–.37
Control	.24	.04	.09	.06–.42

Note. MisRp+, MisRp−, and MisNrp is where misinformation had been presented about an Rp+ (practiced item), Rp− (unpracticed item), and Nrp item (unpracticed item from unpracticed category), respectively. The control condition is where no relevant retrieval practice had taken place. Overall performance on the forced-choice recognition test is calculated on the basis of responses to all items with the exception of the critical misinformation item. Misinfo. = misinformation.

This study indicates that the retrieval practice of a subset of noncritical items can create conditions ideal for the promotion of misinformation effects. By conducting retrieval practice on a subset of items about an event (in this case, items stolen from a house), other items known about that event are actively inhibited. When misleading information is then introduced about these inhibited items, participants are more likely to choose the misleading information during a subsequent forced-choice test of memory. The fact that the level of misinformation reported was significantly higher where misleading information had been introduced about inhibited (i.e., MisRp− condition) than noninhibited items (i.e., MisRp+, MisNrp, and control conditions), suggests that retrieval-induced forgetting has a major role to play in explaining misinformation effects.

Given this potential, we posed a more stringent test of our hypothesis. Specifically, as retrieval-induced forgetting is considered to be the result of the active inhibition or suppression of related items in memory (M. C. Anderson, 1994; M. C. Anderson & Green, 2001; M. C. Anderson & Spellman, 1995); we could expect misinformation effects to be evident only where the critical items remain inhibited. For many kinds of information, this inhibitory mechanism must be transient for it to perform an adaptive role, that is, there is no apparent adaptive advantage in rendering certain kinds of information permanently unavailable given constantly changing goal states (MacLeod & Macrae, 2001; Macrae & MacLeod, 1999). In their examination of some of the boundary conditions of retrieval-induced forgetting, MacLeod and Macrae (2001) recently demonstrated that a critical determinant of this kind of temporary forgetting is the interval between guided retrieval practice and final recall test. When these two phases are separated by a 24-hr interval, retrieval-induced forgetting fails to emerge. When they occur in the same testing session, however, retrieval practice prompts the inhibition of related items in memory. Thus, if misinformation effects are dependent on the inhibition of critical items, such effects should only be evident where inhibition remains active. Specifically, we could expect that a 24-hr delay between retrieval practice and recall to produce neither retrieval-induced forgetting nor misinformation effects. In contrast, we could expect that a 24-hr delay between original study and retrieval practice (i.e., no such delay between retrieval practice and test) to result in both retrieval-induced forgetting and misinformation effects. In other words, misinformation effects would be entirely dependent on the critical items remaining inhibited. We tested these predictions in our second study.

Study 2: The Effects of Delayed Testing and Delayed Practice on Misinformation Effects

Method

Participants and Design

Ninety undergraduate students (37 men and 53 women) participated on a voluntary basis in this study. The experiment had a single factor (timing of delay: delayed test, delayed retrieval practice, or no delay) between-subjects design. Each condition included 30 participants. None of the participants had taken part in the previous study. The delayed-test condition comprised presentation, retrieval practice, 24-hr delay, and final test, whereas the delayed-retrieval practice comprised presentation, 24-hr delay, retrieval practice, and final test. These components also made up the no-delay condition but with the excision of the delay manipulation. Misinformation was introduced before the final test phase in each condition (see Figure 2 for a summary of the experimental procedure).

Procedure and Stimulus Materials

Participants arrived at the laboratory individually or in groups of up to four, were greeted by a female experimenter, and were randomly assigned to one of the test conditions. Instructions to participants, retrieval practice procedures, filler tasks, and test materials were identical to those used in Study 1. As in the previous study, no participants completed any of the filler tasks within the time permitted. Unlike the previous study, however, postevent misinformation was introduced about Rp− items only (i.e., inhibited items). The effect on memory of introducing misinformation about the stolen items was measured with the forced-choice recognition task used in the previous study. On completion, participants were thanked and debriefed before leaving the laboratory.

Results and Discussion

The retrieval practice success rates were 86.7%, 88.4%, and 88.2% in the delayed test, delayed retrieval practice, and

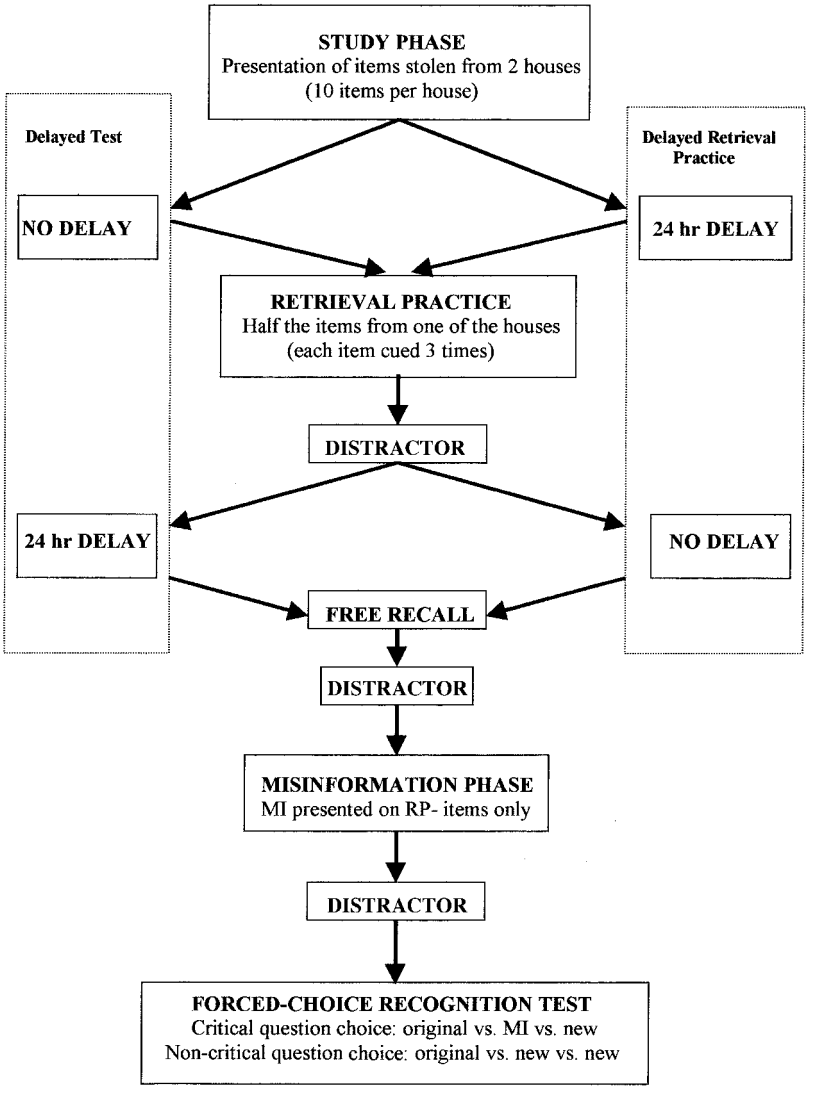

Figure 2. Procedure for Experiment 2.

no-delay conditions, respectively. Manipulation checks were undertaken before the introduction of misleading postevent information. Table 3 indicates that the difference in recall performance for Rp− items (unpracticed items from the practiced set) compared with Nrp items (unpracticed items from the unpracticed set) was −.13 for the no-delay condition and −.14 for the delayed retrieval practice condition—consistent with the magnitude of the retrieval-induced forgetting effects reported in the previous study and other published studies on retrieval-induced forgetting. Our manipulation checks confirmed the presence of the predicted facilitation and retrieval-induced forgetting effects. Single factor (item type: Rp+, or Rp−, or Nrp) within-subjects ANOVAs revealed main effects of item type, $F(2, 58) = 30.41$, $p < .001$, $MSE = 0.048$, Cohen's $f = 1.02$; and, $F(2, 58) = 50.86$, $p < .001$, $MSE = 0.036$, Cohen's $f = 1.32$, for both the delayed retrieval practice and no-delay conditions, respectively. The effect sizes were large for both conditions (Cohen, 1988, p. 284–285).

As in the previous study, we controlled for familywise error rate in a series of post hoc paired-samples *t* tests by using Holm's sequential Bonferroni approach. Participants reported significantly more Rp+ items than Nrp items in both the delayed retrieval practice and no-delay conditions, $t(29) = 5.84$, $p < .001$; and $t(29) = 7.93$, $p < .001$, respectively (i.e., the presence of facilitation effects). Consistent with predictions, paired-samples *t* tests also revealed that Rp− item recall performance was significantly lower than recall performance for Nrp items, $t(29) = -2.38$, $p < .05$; and, $t(29) = -2.55$, $p < .05$, in both the delayed retrieval practice and no-delay conditions, respectively, thereby confirming that retrieval-induced forgetting had taken place. Also, significantly more Rp+ items were recalled than Rp− items, $t(29) = 7.42$, $p < .001$; and, $t(29) = 9.20$, $p < .001$, in both the delayed retrieval practice and no-delay conditions, respectively. See Table 3 for means.

Consistent with our predictions, a single factor (item type: Rp+, or Rp−, or Nrp) within-subjects ANOVA failed to produce a significant effect of item type on recall performance in the delayed test condition, $F(2, 58) = 0.44$, *ns*, thereby confirming the recent findings by MacLeod and Macrae (2001). We realize, however, that there are often difficulties in interpreting data that affirm the null hypothesis, but in this case we can be reasonably confident that the affirmation of the null hypothesis is unlikely to be due to lack of statistical power. Indeed, the power of the analysis is such that the difference between Rp− and Nrp item recall performance ([Rp−] − Nrp) would need to be at least −.096 to have proven significant ($\alpha = .05$). Here, the reported difference was not only smaller but in the opposite direction (+.04, see Table 3).

Table 3
Mean Recall Performance as a Function of Item Type—Study 2

	Item type			Difference score	
Condition	Rp+	Rp−	Nrp	Rp+ − Nrp	Rp− − Nrp
Delayed test	.33 (.22)	.33 (.21)	.29 (.17)	.04	.04
Delayed-retrieval-practice	.80 (.17)	.37 (.26)	.51 (.18)	.29	−.14
No delay	.85 (.14)	.37 (.26)	.50 (.18)	.35	−.13

Note. Standard deviations are enclosed in parentheses. Rp+ items were practiced items from the practiced set, Rp− items were unpracticed items from the practiced set, and Nrp items were unpracticed items from the unpracticed set.

We also conducted a further manipulation check involving an additional 30 participants, which confirmed that the retrieval practice manipulation in the delayed test condition was likely to have produced a significant effect of item type before the delay manipulation, $F(2, 58) = 44.32$, $p < .001$, $MSE = 0.038$, Cohen's $f = 1.24$ (indicating a large effect). Using Holm's sequential Bonferroni approach, post hoc paired-samples *t* tests indicated the presence of both facilitatory, $t(29) = 7.75$, $p < .001$, and retrieval-induced forgetting effects, $t(29) = -2.20$, $p < .05$. Participants also recalled significantly more Rp+ items than Rp− items, $t(29) = 11.58$, $p < .001$. Thus, the predicted lack of retrieval-induced forgetting in the delayed test condition is unlikely to be due to any failure of the retrieval practice procedure to produce retrieval-induced forgetting before the delay manipulation.

The results of these manipulation checks are consistent with the idea that retrieval-induced forgetting effects are transitory (MacLeod & Macrae, 2001). On the introduction of a 24-hr delay between retrieval practice and test, the inhibition of related items set up during the retrieval practice phase dissipates. Once sufficient time has elapsed, these related items are no longer subject to inhibition and are once again available to conscious inspection. In contrast, when a similar delay is introduced between study and retrieval practice, retrieval-induced forgetting remains evident, that is, temporary forgetting emerges even when the study and retrieval practice phases are separated by 24 hr (see Table 3).

Given the differential effects of delay following study and retrieval practice on retrieval-induced forgetting, we now consider their consequences for the report of misinformation. On the basis of our earlier predictions, we predicted that the report of misinformation would be evident only where critical items (i.e., those subject to misinformation) remained inhibited. In contrast, where inhibition has dissipated (e.g., following a 24-hr delay after retrieval practice), we expected to find low levels of reported misinformation. We used the delayed test condition as the control against which the other conditions could be compared for the level of reported misinformation. In fact, a chi-square analysis revealed that there was no difference in the level of reported misinformation between the delayed test in Study 2 and the control condition in Study 1. For both the delayed retrieval practice and no-delay conditions, medium-sized effects were evident where participants were found to be significantly more likely to report misinformation than those in the delayed test condition, $\chi^2(1, N = 60) = 7.05$, $p < .01$, $\Phi = .34$; and, $\chi^2(1, N = 60) = 4.69$, $p < .05$, $\Phi = .28$, respectively. Table 4 indicates that in the delayed retrieval practice and no-delay conditions (i.e., where inhibition had occurred), participants were more than twice as likely to report misinformation for the critical item than where inhibition had dissipated because of the introduction of a 24-hr delay between retrieval practice and final test. Participants in the inhibited conditions were also approximately eight times more likely to make an error on the critical item than on noncritical items. As in the previous study, there was no indication that the report of misinformation was dependent on poor overall recall performance in the final test (94%, 92%, and 94% accuracy for noncritical items in the de-

Table 4
Likelihood of Participants Choosing Correct, Misinformation, and New Erroneous Information During Forced-Choice Recognition Test—Study 2

Condition	Critical item		Noncritical item	95% confidence interval
	Misinfo.	New error	Error	
Delayed test	.20	.10	.06	.05–.35
Delayed-retrieval-practice	.57	.10	.08	.38–.75
No delay	.50	0	.06	.31–.69

Note. Misinformation had been presented about Rp− (unpracticed) items only. Overall performance on the forced-choice recognition test is calculated on the basis of responses to all items with the exception of the critical misinformation item. Misinfo. = misinformation.

layed test, delayed retrieval practice, and no-delay conditions, respectively).

Additional Analyses

Output interference. Although we have argued that our data are consistent with an inhibitory account, we readily acknowledge that alternative explanations could be offered for the present findings. One such possibility is that the observed effects are attributable to the operation of output interference during the final-recall task. In other words, the first items that are produced in a recall task can interfere with the retrieval of related material from memory (Roediger & Schmidt, 1980; Tulving & Arbuckle, 1963). In the present set of studies, we could expect this to take the following form. Given a prior period of retrieval practice, it is possible that participants may initially retrieve highly accessible Rp+ items (i.e., practiced items) from memory. If this were the case, we could expect the retrieval of Rp+ items to interfere with the retrieval of less accessible Rp− items (i.e., unpracticed items), thereby diminishing recall performance for these items. If output interference was operating in the present set of studies, we could expect a number of predictable effects. Specifically, the pattern of results indicative of retrieval-induced forgetting should only emerge where participants retrieve Rp+ items early in the memory task.

To establish this possibility, we followed the statistical procedure outlined by Macrae and MacLeod (1999). For each of the reported experimental conditions where retrieval-induced forgetting had been evident, we classified participants according to the extent to which they commenced their recall sequences with Rp+ or Rp− items. This was achieved by giving each participant a score that represented the extent to which the recall sequence began with Rp+ or Rp− items. This was a difference score that was calculated by subtracting the average recall position of Rp+ items from the average recall position of Rp− items (i.e., negative values reflect early Rp− output, positive values reflect early Rp+ output). Participants in each counterbalanced group (i.e., those who performed retrieval practice on items from the Smith's house vs. those who performed retrieval practice on items from the Jones' house) were then sorted by a median split on these scores. The bottom halves of each group were then combined to form the early Rp− group and the top halves combined to form the early Rp+ group. The inhibition effects for these two groups were then calculated in the normal way (i.e., [Rp−] − Nrp). It is noteworthy that in none of the reported conditions did the early Rp+ group produce a significantly larger inhibitory effect than the early Rp− group: Study 1, MisRp+ condition (Ms = −0.13 vs. −0.15), $t(23) = 0.11$, *ns;* MisRp− condition (Ms = −0.15 vs. −0.19), $t(23) = 0.46$, *ns;* MisNrp condition (Ms = −0.14 vs. −0.24), $t(23) = 1.02$, *ns;* Study 2, delayed retrieval practice condition, (Ms = −0.18 vs. −0.13), $t(28) = -0.38$, *ns;* no-delay condition (Ms = −0.13 vs. −0.13), $t(28) = 0.06$, *ns.*

Although these analyses eliminate output interference as a plausible explanation for the retrieval-induced forgetting effects observed in the present studies, the data do not in themselves represent incontrovertible evidence that inhibition underlies these effects. Nevertheless, the data are consistent with patterns observed in other studies where evidence for inhibitory mechanisms is arguably stronger (see M. C. Anderson & Green, 2001; M. C. Anderson & Neely, 1996; M. C. Anderson & Spellman, 1995). One of the central premises of inhibitory accounts of memory is that if a representation is truly inhibited, then impairment in recall performance arising from that inhibition should generalize to any cue used to test that item. In other words, forgetting should be cue-independent (M. C. Anderson & Spellman, 1995). This contrasts markedly with noninhibitory accounts of memory where impairment in recall performance is attributed to changes in cue-target associations (cf. Tulving, 1974). If changes in cue-target associations are responsible for forgetting an item, then impairment should disappear only once a distinct retrieval cue is used (e.g., the impairment of *whisky* caused by retrieval practice of *drink–sherry* should disappear when *whisky* is tested with *Scotsman*). Thus, cue independence serves as a general empirical criterion by which inhibition in memory can be established. Taken together, the combination of the present studies and published work on cue independence provides converging evidence for the operation of inhibitory mechanisms in retrieval-induced forgetting.

Cross-category inhibition. A second tenet of inhibition theories of memory is that retrieval practice should result not only in the inhibition of unpracticed items in the practiced set but also in related items in the unpracticed set. Consider two sets of items, one of *food* (*food–crackers, food–strawberry*), another of *red* (*red–blood, red–tomato*). Under conditions of this type, retrieval practice of *red–blood* impairs the recall of *red–tomato* and the recall of *food–strawberry* but not *food–crackers* (see M. C. Anderson & Spellman, 1995). Noninhibitory theories of memory would predict that the recall of the item *strawberry* should not be impaired by the retrieval practice procedure, whereas inhibitory accounts suggest that such an effect is possible because of the implicit link between *red* and *strawberry,* leading to the inhibition of *strawberry.* If retrieval-induced forgetting involves inhibitory processes acting at the representational level, then retrieval-induced forgetting should be observed in all items that are related to the inhibited items (whether by initial set or other semantic links). Such inhibitory mechanisms have been demonstrated empirically by M. C. Anderson and Spellman (1995), and more recently by M. C. Anderson and Green (2001).

Although the typical finding in the literature is that recall performance for Nrp items is unaffected by the retrieval practice of Rp+ items in comparison with a between-subjects control baseline (see, e.g., MacLeod, 2002; Macrae & MacLeod, 1999; Shaw et al.,

1995), the fact that there is significant impairment in the level of Nrp items recalled in comparison with this baseline control (Study 1) may indicate the operation of cross-category inhibition (see Table 1). In the present studies, this form of inhibition is possible given that both sets of items are highly related to each other. Not only are the items in each set described as being "stolen items from a household," but they are also similar in a more basic sense, that is, there is reasonable degree of semantic correspondence between items in the two sets. For example, the item *printer* in the Smith's house is semantically linked to the item *computer* in the Jones' house. Similarly, the item *television* in the Smith's house is closely associated with the item *video recorder* in the Jones' house. Although there is no direct one-to-one correspondence between all items in both houses, there is a pattern of relatedness that has arisen inadvertently as a result of generating experimental materials about "stealable" things. Thus, it is plausible that the act of practicing items from one household could cause inhibition not only of nonpracticed items in the same set but also related unpracticed items from the unpracticed set, even though these items were studied under different household cues.

Following this line of logic, if cross-category inhibition is operating in the present studies, an inhibitory account of the misinformation effect becomes less tenable as we could also expect a high level of misinformation to be reported in the MisNrp condition. Specifically, we could predict that the level of misinformation in the MisNrp condition to be somewhere between that obtained for the MisRp+ condition (i.e., where misinformation had been introduced about noninhibited items) and the MisRp− condition (i.e., where misinformation had been introduced about inhibited items). We would not expect the level of misinformation reported in the MisNrp condition to be at a comparable level to that obtained in the MisRp− condition because of the randomization procedure used in choosing items in the retrieval practice phase for each participant. Assuming that cross-category inhibition has occurred, the randomization procedure would have ensured that misinformation is introduced about an inhibited Nrp item for only a proportion of the participants in the MisNrp condition. The remaining participants in this condition would have received misinformation about a noninhibited Nrp item. The fact that there was neither a significant difference in the level of misinformation reported between the MisRp+ and MisNrp conditions (*M*s = .16 vs. .20), nor between the MisNrp and control conditions (where no relevant retrieval practice had taken place and therefore no possibility of cross-category inhibition), strongly suggests that cross-category inhibition had not occurred (*M*s = .20 vs. .24; see Table 2).

To confirm this, we performed a series of paired-samples *t* tests in which recall performance for Nrp-similar items (i.e., Nrp items semantically related to Rp+ items) was compared with recall performance for Nrp-dissimilar items (i.e., Nrp items semantically unrelated to Rp+ items). The mean recall performance for Nrp-similar versus Nrp-dissimilar was as follows for each condition: Study 1; MisRp+ condition (*M*s = 0.35 vs. 0.37), $t(24) = -0.31$, *ns;* MisRp− condition (*M*s = 0.44 vs. 0.36), $t(24) = 1.10$, *ns;* MisNrp condition (*M*s = 0.47 vs. 0.39), $t(24) = 0.95$, *ns;* Study 2; delayed retrieval practice condition (*M*s = 0.48 vs. 0.49), $t(29) = -0.15$, *ns;* no-delay condition (*M*s = 0.49 vs. 0.38), $t(29) = 2.22$, $p < .05$. We also conducted a parallel set of paired-samples *t* tests in which we considered the level of recall performance for Nrp items that were either similar or dissimilar to Rp− items. However, as in the previous analysis, there was no evidence that Nrp-similar items had been recalled more poorly than Nrp-dissimilar items. In other words, there is no evidence of cross-category inhibition having taken place in either of the studies reported in the present article. Thus, the drop in Nrp performance relative to controls in Study 1 (Table 1) does not pose a threat to our thesis that inhibition facilitates the report of misinformation. It should also be noted that although we found no evidence of cross-category inhibition in the present studies, this does not constitute evidence that inhibitory mechanisms are not responsible for the observed retrieval-induced forgetting effects. Indeed, one of the reasons for the failure to show such cross-category effects may be because of the issue of nominal versus functional similarity (M. C. Anderson & Spellman, 1995). In other words, although as experimenters we might see the semantic relatedness of items such as *Discman* and *mobile phone*, unless there are enough of these items on the list to make these similarities salient, participants are unlikely to have encoded such relationships. Thus, as our studies were not designed a priori to consider cross-category inhibition, it remains for future studies to design more optimal tests by which the impact of cross-category inhibition on the likelihood of reporting misinformation can be explored.

Source confusion. The absence of cross-category inhibition, however, raises the intriguing issue as to whether noninhibitory mechanisms may have contributed to the decrease in Nrp recall performance. One possibility is that it is due to source confusion. As noted earlier, the two sets of items are highly similar to one another. In addition, they are grouped together into essentially arbitrary sets (i.e., the items in Smith's and Jones' households have no intrinsic meaning). Given the substantial delay (24 hr) in the delayed retrieval practice condition (Study 2), it would seem plausible that participants could have suffered source confusion about which items were studied under which household. Thus, when recalling items from the Smith's household, participants may have recalled items from the Jones' household and vice versa. Clearly, if these had been scored as incorrect, this source confusion could account for the low Nrp baseline and also perhaps the low Rp+ recall performance in the delayed test condition in Study 2. Other studies that have used similar delays in examining retrieval-induced forgetting effects (MacLeod & Macrae, 2001) did not show such a marked drop in Rp+ performance. More important, the improvement in the score for the Nrp baseline in the delayed practice condition could be because of the possibility that delayed retrieval practice may have alleviated such source confusion. For example, the act of practicing the item *television* as having been stolen from the Smith's household could lead participants to put the item *video recorder* under the Jones' household by a process of elimination (assuming that participants had some metaknowledge that there were similar confusable items in both sets). However, the items recalled in the present studies were coded as being correct irrespective of whether they had been associated with the correct household, thereby eliminating source confusion as a possible contributor to the pattern of data obtained.

Category dropout. Nevertheless, the low Nrp performance observed in the delayed test condition (Study 2) may still reflect the effects of interhousehold similarity. Basically, it could be argued that the items listed under the two households could be construed as falling into two subcategories that span the house-

holds (e.g., electrical vs. nonelectrical goods). It has been known for some time that in the free recall of categorized lists, recall can be poor for two reasons: the forgetting of categories, and the forgetting of items within categories. Sometimes people simply forget whole categories of things, but when they are reminded of that category, they can instantly recall large numbers of members from that category. Thus, free recall of a categorized list is best thought of as a two-stage process. The first involves the recall of the category, given the context as cue, followed by the recall of items within that category, given the category and context as cues (see Rundus, 1971; Tulving & Pearlstone, 1966).

As subcategories span both households, the two sets are potentially interdependent. Thus, retrieval practice on some items from the practiced category following a 24-hr delay may lead to an increase in the overall accessibility of that implicit category (e.g., electrical goods). This may increase the likelihood that items from that subcategory would be recalled when recalling the Nrp household items. This, in turn, could account for the boost in recall of Nrp items in the delayed-practice condition (see Table 3). By this account, the within-category inhibition has not gone away in the delayed-test condition, but rather that the subcategories have dropped out and need to be cued in order to reveal Rp− impairment. In the present studies, however, care was taken to prevent Rp+ items constituting such subcategories (i.e., Rp+ sets comprised both electrical and nonelectrical items and varied from participant to participant in terms of the actual items used). Previous research has shown that where coherent subcategories constitute Rp+ and Rp− sets, retrieval-induced forgetting fails to emerge, presumably because the items in the Rp− set no longer provide competition for retrieval with items in the Rp+ set (see Oram & MacLeod, 2001).

Clearly, we cannot entirely eliminate the possibility that some form of category dropout had occurred in the present studies. For example, participants may have been more likely to experience category dropout for *nonelectrical* items given that such items represent a less coherent set than do items that constitute an *electrical* set. To test this possibility, we computed the proportion of Nrp items recalled for each category (i.e., electrical vs. nonelectrical) for each of the conditions where retrieval-induced forgetting was observed: Study 1; MisRp+ condition (Ms = 0.38 vs. 0.42), $t(24) = -0.71$, *ns;* MisRp− condition (Ms = 0.41 vs. 0.42), $t(24) = -0.21$, *ns;* MisNrp condition (Ms = 0.41 vs. 0.46), $t(24) = -0.74$, *ns;* Study 2; delayed retrieval practice condition (Ms = 0.53 vs. 0.47), $t(29) = 1.09$, *ns;* no-delay condition (Ms = 0.42 vs. 0.51), $t(29) = -1.92$, *ns*. This analysis indicates that there is no evidence of category dropout (either electrical or nonelectrical) in either study. Although it remains possible that participants may have chosen different categories to those used in the present analysis, the randomization procedure used in the selection of Rp+ items and the relatively small size of remaining subcategories makes it unlikely that further analysis would demonstrate category dropout.

General Discussion

Recognizing the fact that the misinformation effect is a well-established psychological phenomenon but one that has garnered little consensus regarding its interpretation (Wright & Davies, 1999), the present article examined the possibility that retrieval-induced forgetting may provide a mechanism by which misinformation effects occur. In particular, the present article examined whether retrieval-induced forgetting could explain the observation that, under certain circumstances, memory for original and postevent information coexist but, under other conditions, only memory for postevent material appears to be available to conscious inspection. Specifically, in our first study we considered whether the inhibition of critical items (i.e., as a consequence of the action of retrieval practice on a subset of related items) would render the recollection of postevent information more likely in a subsequent test of memory. In our second study, we considered the relationship between the transient nature of retrieval-induced forgetting (MacLeod & Macrae, 2001) and the presence of misinformation effects. Under conditions where retrieval-induced forgetting is known to dissipate, would misinformation effects remain evident?

Study 1 established that significantly more errors were reported where postevent misinformation had been introduced about inhibited (i.e., Rp− items) than noninhibited items (i.e., Rp+, or Nrp, or control). Participants in the MisRp− condition were almost six times more likely to make a mistake on the critical item than on noncritical items (see Table 2). It is unlikely that this effect can be explained by the operation of demand characteristics (cf. McCloskey & Zaragoza, 1985), as we could have expected misinformation effects to exist across all conditions given that any demand characteristics would have been equivalent. Similarly, it is difficult to see how a failure to encode the critical item or some form of "spontaneous" forgetting could account for the observed pattern of data. If such factors had contributed in any significant way, we could have expected a significant number of participants to have chosen the completely new erroneous information by chance alone in conditions where the critical item had not been subject to retrieval-induced forgetting. Our data, however, strongly indicate that misinformation effects were present only in those conditions where retrieval-induced forgetting had occurred for critical items. In other words, there is strong inference that retrieval-induced forgetting can significantly boost the likelihood of misinformation being reported where it has been introduced about items that have inhibited as a result of the retrieval practice of a subset of related items.

Our interpretation gains further support from our second study where misinformation effects were evident only where inhibition remains active. Specifically, we found that where misinformation had been presented about Rp− items, a 24-hr delay between retrieval practice and final test resulted in no retrieval-induced forgetting, and the elimination of the typical misinformation effect (as observed in the MisRp− condition in Study 1). In contrast, where there was no delay or a delay of 24 hr between original presentation and retrieval practice, both retrieval-induced forgetting and misinformation effects were observed, comparable to those reported in the MisRp− condition (Study 1). Also, the observed pattern of retrieval-induced forgetting following such delays replicates recent findings by MacLeod and Macrae (2001).

Our studies set out to determine whether retrieval-induced forgetting could provide the conditions necessary to promote misinformation effects. The data obtained in the two studies reported herein support this relationship. Inevitably, however, such research raises further questions of theoretical and applied interest. In particular, the present studies do not reveal whether it is the

inhibition of the critical item at time of final test or the inhibition of the critical item during the introduction of misinformation that is essential for the production of misinformation effects. In the delayed test condition, the critical item was no longer inhibited because of the introduction of a 24-hr delay at the time of both final test and the introduction of the misinformation. In the delayed retrieval practice condition, in contrast, the critical item was inhibited at the time of both final test and the introduction of misinformation. Of course, it may be the case that inhibition needs to be active at both points in time (as illustrated by the present studies) in order to facilitate the production of misinformation effects.

This point is of some importance if we are to try to dissociate between retrieval-induced forgetting accounts of the misinformation effect and other viable competing explanations such as the decay of the memory trace over time. Loftus and colleagues (1978), for example, argued that the introduction of misinformation following a delay gives rise to a strong misinformation effect because it is easier to alter original memory where the memory trace has been weakened. If the decay of the memory trace had been a significant factor in the present studies, however, we could have expected a difference in recall performance for noncritical items between the delay and no-delay conditions (Study 2). However, there was no such difference. Rather, the only difference concerned accuracy of memory for the critical item. Where we had predicted that the critical item would be inhibited, the misinformation effect was evident; and where we had predicted no inhibition of the critical item, no misinformation effect was produced. In other words, for this particular paradigm and set of materials, it would appear that the presence of a misinformation effect is due to an item-specific mechanism (i.e., the forgetting of unpracticed items from the practiced set) rather than a more general loss of memory performance for original material. Of course it remains possible that under other conditions, retrieval-induced forgetting could interact with memory trace decay. Indeed, it is conceivable that, as memory for an event decays over time, retrieval-induced forgetting effects could be produced even more readily. It should also be noted that in order for retrieval-induced forgetting to take place, retrieval practice need not occur immediately after learning the original material. Recent studies have demonstrated such effects for previously learned material (Koutstaal, Schacter, Johnson, & Galluccio, 1999; MacLeod & Macrae, 2001). Thus, in a forensic setting, it is possible that retrieval-induced forgetting and the introduction of misinformation could occur well after the initial police investigation. The extent to which the decay of the original memory trace facilitates such effects remains to be determined.

Throughout this article, our account of the relationship between retrieval-induced forgetting and misinformation effects has been an inhibitory one (cf. M. C. Anderson & Green, 2001; M. C. Anderson & Spellman, 1995) and, as such, is consistent with accounts of the ubiquity of inhibitory processes in mental life (see Conway, 2001). We consider the most credible explanation of retrieval-induced forgetting to result from the inhibition or suppression of critical items (i.e., Rp− items) as a result of the retrieval practice of other related items (i.e., Rp+ items). In saying this, however, we readily acknowledge that noninhibitory accounts can also contribute to the production of retrieval-induced forgetting effects (see J. R. Anderson, 1983; Bauml, 1998; Raajimakers & Shiffrin, 1981). In the present article, our analyses revealed that neither output interference, category dropout, nor source confusion significantly contributed to the observed effects. We also believe that the weight of available empirical evidence strongly favors an inhibitory account (e.g., the independent probe technique; M. C. Anderson & Spellman, 1995), cross-category forgetting (M. C. Anderson & Green, 2001; M. C. Anderson & Spellman, 1995); and the fact that repeated presentation rather than recall of a subset of items is insufficient to produce a decrement in memory for related material (M. C. Anderson et al., 2000).

It is important to note, however, that the fact that our second study is consistent with the idea that retrieval-induced forgetting effects are transient (MacLeod & Macrae, 2001), the temporary nature of retrieval-induced forgetting is not in itself intrinsic to inhibitory theory (see M. C. Anderson & Spellman, 1995, for a discussion). Although we have argued elsewhere (MacLeod & Macrae, 2001) that there is little adaptive value in forgetting information that one might need at some later point in time, there are probably exceptions where retrieval-induced forgetting is of a more permanent nature. M. C. Anderson and Spellman (1995) made the cogent point that the forgetting of inhibited material for long periods of time need not necessarily be maladaptive. Thus, it is possible that the forgetting of many one-time incidental facts may actually serve to reduce confusion and clutter. Although we could expect the long-term forgetting of highly used knowledge to be maladaptive, there is also long-term knowledge that offers little value in terms of retention (e.g., the phone numbers of past girlfriends, details of the addresses of past residences). Indeed, for most everyone, the retention of such facts in memory is not only very difficult but serves little obvious purpose. Thus, it is important not to overstate the view that long-term forgetting is dysfunctional. Indeed, it remains possible that retrieval-induced forgetting effects may be long-lasting and, where this occurs, should neither be taken as evidence against an inhibitory account or that it is necessarily maladaptive.

It is also important to acknowledge that although the only strong misinformation effects in the present studies occurred where critical items had been inhibited, it remains possible that other routes are capable of producing misinformation effects. For example, the present studies do not deny the possibility that, for short duration experiments where the source element of memory has not decayed, misled participants may still report misinformation because of source misattribution or because they simply did not attempt to remember the source of the content memory. Notwithstanding such possibilities, the current research provides a valuable new insight into a possible mechanism by which misinformation effects occur. In doing so, this research sheds new light not only on our understanding of how misinformation can be incorporated into witness reports, but also a mechanism by which misinformation effects are produced in the standard misinformation paradigm (e.g., Loftus et al., 1978). There are, however, a number of differences between the procedures adopted in the current set of studies and those used by Loftus and colleagues (e.g., misinformation in the standard paradigm is introduced at the same time as when retrieval practice takes place). To determine that retrieval-induced forgetting had occurred in the present studies, misinformation was introduced subsequent to the retrieval practice procedure. Despite such differences, we remain confident that our studies illuminate a possible mechanism by which misinformation effects are produced in the standard paradigm. We reemphasize, however, that far from

detracting from the pioneering work of Loftus and colleagues, the current research provides novel and additional support for the relevance of their work to our understanding of real-world problems. In particular, our findings demonstrate how it is possible for misinformation to be inadvertently introduced as part of the investigative procedures that typically follow a crime or accident. Clearly, the more that we understand about the mechanisms underlying misinformation effects, the better placed we would be to anticipate and minimize their effects.

Consistent with this endeavor, there are a number of additional characteristics in the data that may shed further light on the underlying mechanism responsible for the misinformation effects reported in the present article. In particular, what kind of mechanism could account for the low Nrp within-subjects baseline performance in the experimental conditions in comparison with the between-subjects baseline in Study 1. Similarly, what could account for the low Nrp baseline in the delayed test condition with the other conditions in Study 2? Given that our additional analyses eliminated cross-category inhibition, and other noninhibitory mechanisms such as source confusion and category dropout as potential explanations for the pattern of data obtained, what might be the reason? One possibility derives from a model we have recently developed to explore the mechanisms underlying retrieval-induced forgetting (Oram & MacLeod, 2001). It is important to note that the following does not constitute a validation of our model, as we did not aim to test the model in the current set of studies. Rather, the experimental data are consistent with what the model would predict and therefore of interest in terms of its future potential power in predicting misinformation effects.

Our model uses a modified Hebbian learning algorithm (Oram & Foldiak, 1996) in which we argue that inhibition operates at the level of association nodes. As a result of this inhibition, category representations can be kept distinct while the mechanism underlying retrieval-induced forgetting operates in terms of blocking the associations between items and cue. The learning rule is such that it would enhance the connection strength between two active units and decrease the connection strength between active and inactive units. Given this set up, we could imagine an output unit representing the (partial) association of two input units with their category membership. If only one of those units is practiced, there would be activation of one item-input node and the association node, whereas the second item input remains inactive. This implies that selective practice would enhance the representation of the practiced items and decrease the representation of unpracticed items, thereby giving rise to retrieval-induced forgetting.

Inhibitory processes between association nodes (categories) ensure that only a single category node is active at any one time. A decrease in representational strength would therefore only occur for those items that are connected to active association nodes, leaving the representation of items in unrelated (inactive) categories unchanged. If we consider cue independence—a central feature of inhibitory accounts of retrieval-induced forgetting (M. C. Anderson & Green, 2001; M. C. Anderson & Spellman, 1995)—these can be viewed as a process in which both the initial presentations of the category-exemplar pairs establish a cue-independent associative block and the retrieval practice phase causes first-order inhibitory effects, also by associative blocking. Cross-category inhibition, however, involves the spreading of activity through the system during both the initial learning and retrieval practice phases. It is important to note, however, that inhibition between the units, which maintain category specificity, is still vital for both cue independence and cross-category inhibition to occur. Thus, this model is not suggesting that inhibition provides an inappropriate account of retrieval-induced forgetting but rather tries to specify the levels at which these inhibitory mechanisms are operating.

According to our model, therefore, retrieval practice would produce retrieval-induced forgetting at the category level, which, in turn, would depress Nrp recall performance relative to a between-subjects baseline control (as demonstrated in the MisRp+, MisRp−, and MisNrp conditions; Study 1). Where nonrelevant retrieval practice has taken place (i.e., control condition, Study 1), two levels of sets are established: one concerning the items presented in the study phase of the experiment (i.e., items from both households), and the other being concerned with capital cities of the world (i.e., nonrelevant retrieval practice). The former set does not provide retrieval competition with the latter, as no cross-links exist between items in the two sets and therefore no partial activation would occur. Relevant retrieval practice, however, would significantly strengthen links between the category and Rp+ items and, concurrently, weaken links between the category and Nrp items. Thus, our model would predict the observed pattern of data for Study 1; that is, that the within-subjects Nrp baseline in the experimental conditions should be lower than the between-subjects Nrp baseline in the control condition.

Our model also provides a post hoc explanation for the pattern of data obtained in Study 2. Specifically, participants in the delayed test conditions would have formed and practiced other memories during the 24-hr delay between retrieval practice and final test. Some of these other memories for events, objects, and people would have been practiced more than the information originally learned about the items stolen from the Smith's and Jones' households. This would result in the suppression of the learning node that, in turn, would result in the general suppression of both the Rp and Nrp sets. Thus, in the delayed test condition, we might expect this to give rise to the rather flat recall performance that was obtained across all item types. In the delayed retrieval practice condition, the same general suppression of the learning node would have occurred as a result of repeatedly retrieving other memories. However, because the retrieval practice of items in the Rp+ set occurs after this delay, there would be an enhancement in recall performance for Rp+ items and a consequent drop in recall performance for Rp− items (i.e., retrieval-induced forgetting). The Nrp items would remain unaffected by the retrieval practice and would therefore remain at the level caused by the general suppression of the learning node. The pattern of data obtained in the no-delay condition (Study 2) can be explained in exactly the same way as that for the MisRp− condition (Study 1) given that it is effectively a replication. It should be noted that although our model provides a post hoc rationale for the observed data based on multiple levels of representation and the consequences that these have for the levels at which inhibition occurs, it fails to explain the difference in Nrp recall performance across the two experiments. This may simply indicate the operation of other mechanisms that have gone undetected by our analyses. Future studies, however, would need to pay attention to the consequences of retrieval practice on Nrp recall performance and consider how these relationships might be best explored.

In conclusion, the present research extends our knowledge of how retrieval-induced forgetting can pervade everyday life. More important, such research not only provides a novel interpretation of misinformation effects, but also has the potential to reveal new insights into our understanding of eyewitness reliability, false memory syndrome, recovery from trauma, traumatic amnesia, the design of effective interview and therapeutic practices, as well as the development of theoretical models of memory. In short, it is conceivable that any procedure that involves the asking of questions about another person or event could generate exactly those kinds of conditions that, in turn, promote misinformation effects and consequent misunderstandings. Such situations range from the employment interview, to the interrogation of suspects and witnesses to crimes, to the formation of impressions about other individuals and stereotypes in society, to how autobiographical information is elicited in clinical and forensic settings. In each case, the common denominator is that the information requested is likely to represent a subset of the information known about that particular topic, event, or person. It is also possible that self-generated goals could give rise to retrieval-induced forgetting. For example, it is conceivable that retrieval-induced forgetting may be implicated in the maintenance of self-image. To what extent does our tendency to recall the good things that we have done actively inhibit our memory for the not-so-good things? To what extent are such memorial processes related to the maintenance of self-esteem? Conversely, to what extent do depressed individuals inhibit memories for positive information about themselves by retrieving memories for negative information? Similarly, this field of research may prove to have important ramifications for study techniques. Is it better to do last-minute cramming for exams and risk inhibiting memory for other related material or take the chance that the selection of hastily revised topics may appear in the exam? Clearly, there is much to be accomplished in further delineating the boundary conditions of retrieval-induced forgetting. The present studies, however, confirm the potential of such research not only for our understanding of applied problems, such as the fallibility of memory, but also future theorizing about memory function. In time, this line of inquiry may lead us to a better appreciation of what William James meant by the interdependence of forgetting and remembering.

References

Anderson, J. R. (1983). *The architecture of cognition.* Cambridge, MA: Harvard University Press.

Anderson, M. C., Bjork, E. L., & Bjork, R. A. (2000). Retrieval-induced forgetting: Evidence for a recall-specific mechanism. *Psychonomic Bulletin & Review, 7,* 522–530.

Anderson, M. C., Bjork, R. A., & Bjork, E. L. (1994). Remembering can cause forgetting: Retrieval dynamics in long-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20,* 1063–1087.

Anderson, M. C., & Green, C. (2001, March 15). Suppressing unwanted memories by executive control. *Nature, 410,* 366–369.

Anderson, M. C., & McCulloch, K. C. (1999). Integration as a general boundary condition on retrieval-induced forgetting. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25,* 608–629.

Anderson, M. C., & Spellman, B. A. (1995). On the status of inhibitory mechanisms in cognition: Memory retrieval as a model case. *Psychological Review, 102,* 68–100.

Bauml, K. (1998). Strong items get suppressed, weak items do not: The role of item strength in output interference. *Psychonomic Bulletin & Review, 5,* 459–463.

Bekerian, D. A., & Bowers, J. M. (1983). Eyewitness testimony: Were we misled? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 9,* 139–145.

Belli, R. F. (1988). Color blend retrievals: Compromise memories of deliberate or deliberate compromise responses? *Memory & Cognition, 16,* 314–326.

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

Bjork, E. L., Bjork, R. A., & MacLeod, M. D. (2002, March). *Types and consequences of forgetting: Intended and unintended.* Paper presented at the Third Tsukuba International Conference on Memory: Memory and Society, Tsukuba, Japan.

Bjork, R. A. (1989). Retrieval inhibition as an adaptive mechanism in human memory. In H. L. Roediger & F. I. M. Craik (Eds.), *Varieties of memory and consciousness: Essays in honor of Endel Tulving* (pp. 309–330). Hillsdale, NJ: Erlbaum.

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

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Conway, M. A. (2001, March 15). Repression revisited. *Nature, 410,* 319–320.

Johnson, H. (1994). Processes of successful intentional forgetting. *Psychological Bulletin, 116,* 274–292.

Koutstaal, W., Schacter, D. L., Johnson, M. K., & Galluccio, L. (1999). Facilitation and impairment of event memory produced by photograph review. *Memory & Cognition, 27,* 478–493.

Landauer, T. K., & Bjork, R. A. (1978). Optimum rehearsal patterns and name learning. In M. M. Grunneberg, P. E. Morris, & R. N. Sykes (Eds.), *Practical aspects of memory* (pp. 625–632). London: Academic Press.

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

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

Lindsay, D. S., & Johnson, M. (1989b). The reversed suggestibility effect. *Bulletin of the Psychonomic Society, 27,* 111–113.

Loftus, E. F. (1977). Shifting human color memory. *Memory & Cognition, 5,* 696–699.

Loftus, E. F. (1979). *Eyewitness testimony.* Cambridge, MA: Harvard University Press.

Loftus, E. F., Donders, K., Hoffman, H. G., & Schooler, J. W. (1989). Creating new memories that are quickly accessed and confidently held. *Memory & Cognition, 17,* 607–616.

Loftus, E. F., & Hoffman, H. (1989). Misinformation and memory: The creation of new memories. *Journal of Experimental Psychology: General, 118,* 100–104.

Loftus, E. F., Korf, N., & Schooler, J. W. (1989). Misinformation and memory. In J. Yuille (Ed.), *Credibility assessment: A theoretical and research perspective* (pp. 155–173). Dordrecht, the Netherlands: Kluwer.

Loftus, E. F., & Loftus, G. R. (1980). On the permanence of stored information in the human brain. *American Psychologist, 35,* 409–420.

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 and Memory, 4,* 19–31.

Loftus, E. F., Schooler, J. W., & Wagenaar, W. (1985). The fate of memory: Comment on McCloskey and Zaragoza. *Journal of Experimental Psychology: General, 114,* 375–380.

MacLeod, M. D. (2002). Retrieval-induced forgetting in eyewitness memory: Forgetting as a consequence of remembering. *Applied Cognitive Psychology, 16,* 135–149.

MacLeod, M. D., & Macrae, C. N. (2001). Gone but not forgotten: The transient nature of retrieval-induced forgetting. *Psychological Science, 12,* 148–152.

Macrae, C. N., & MacLeod, M. D. (1999). On recollections lost: When practice makes imperfect. *Journal of Personality and Social Psychology, 77,* 463–473.

McCloskey, M., & Zaragoza, M. (1985). Misleading postevent information and memory for events: Arguments and evidence against memory impairment hypotheses. *Journal of Experimental Psychology: General, 114,* 1–16.

Metcalfe, J. (1990). A Composite Holographic Associative Recall model (CHARM) and blended memories in eyewitness testimony. *Journal of Experimental Psychology: General, 119,* 145–160.

Morton, J., Hammersley, R. H., & Bekerian, D. A. (1985). Headed records: A model for memory and its failures. *Cognition, 10,* 1–23.

Neisser, U. (1981). John Dean's memory: A case study. *Cognition, 9,* 1–22.

Oram, M. W., & Foldiak, P. (1996). Learning generalisations and localisations: Competing for stimulus type and receptive field. *Neurocomputing, 11,* 297–321.

Oram, M. W., & MacLeod, M. D. (2001). Remembering to forget: Modeling inhibitory and competitive mechanisms in human memory. In J. D. Moore & K. Stenning (Eds.), *Proceedings of the Twenty-third Annual Conference of the Cognitive Science Society* (pp. 738–743). Mahwah, NJ: Erlbaum.

Pezdek, K. (2001). A cognitive analysis of the role of suggestibility in explaining memories for abuse. *Journal of Aggression, Maltreatment & Trauma, 4,* 73–85.

Raajimakers, J. G. W., & Shiffrin, R. M. (1981). Search of associative memory. *Psychological Review, 88,* 93–134.

Roediger, H. L., III, & Schmidt, S. R. (1980). Output interference in the recall of categorized and paired associate lists. *Journal of Experimental Psychology: Human Learning and Memory, 3,* 174–188.

Rundus, D. (1971). Analysis of rehearsal processes in free recall. *Journal of Experimental Psychology, 89,* 63–77.

Schooler, J. W., & Tanaka, J. W. (1991). Composites, compromises, and CHARM: What is the evidence for blend memory representations? *Journal of Experimental Psychology: General, 120,* 96–100.

Shaw, J. S., III, Bjork, R. A., & Handal, A. (1995). Retrieval-induced forgetting in an eyewitness-memory paradigm. *Psychonomic Bulletin & Review, 2,* 249–253.

Tulving, E. (1974). Cue-dependent forgetting. *American Scientist, 62,* 74–84.

Tulving, E., & Arbuckle, T. Y. (1963). Input and output interference in short-term associative memory. *Journal of Experimental Psychology, 72,* 89–104.

Tulving, E., & Pearlstone, Z. (1966). Availability versus accessibility of information in memory for words. *Journal of Verbal Learning and Verbal Behavior, 5,* 381–391.

Weinberg, H. I., Wadsworth, J., & Baron, R. S. (1983). Demand and the impact of leading questions on eyewitness testimony. *Memory & Cognition, 11,* 101–104.

Weingardt, K. R., Loftus, E. F., & Lindsay, D. S. (1995). Misinformation revisited: New evidence on the suggestibility of memory. *Memory & Cognition, 23,* 72–82.

Wright, D. B., & Davies, G. M. (1999). Eyewitness testimony. In F. T. Durso, R. S. Nickerson, R. W. Schvaneveldt, S. T. Dumas, D. S. Lindsay, & M. T. I. Chi (Eds.), *Handbook of applied cognition* (pp. 789–818). Chichester, England: Wiley.

Appendix A

Jones' House

The Jones' house had been burgled at New Year. Mr. and Mrs. Jones were out at an office party, while their children were at their friends. Their daughter arrived back at 11 p.m. shortly before her parents. Police believe that the burglars gained access to the Jones' house through the study window situated on the ground floor at the back of the house. The window had been left unlocked. No one saw the burglars break in although neighbors told the police that they had seen a white van parked across the street from the Jones' house before they had gone out to the pub at 9:30 p.m. Police believe that the burglars had moved a picnic table in the back garden to underneath the study window in order to gain access to the house. Footprints in the mud suggested that there were three individuals present. Attached is a list of items that were stolen (underlined).

The binoculars had been in the conservatory at the back of the house. They were used for spotting wildlife in the garden and were next to several nature and wildlife books.

The wristwatch had been upstairs in the master bedroom. It had been lying on the bedside table. The burglars had broken a lamp that was next to it.

The Game Boy had been in the children's room lying on the bed. The burglars never found the games for it, which were on the shelf.

The sunglasses had been in the kitchen next to the kettle and some holiday brochures. They were an expensive designer pair that had been bought last summer.

The television had been in the sitting room which is situated at the front of the house. It was sitting in the corner of the room. The remote control for it hadn't been taken.

The painting had been in the dining room near a table that had art books sitting on it. It had a silver engraved frame. An empty wallet was found near the table.

The printer had been in the study which is at the back of the house. It was on a desk and the burglars had knocked a ream of paper that was next to it all over the floor.

The mobile phone had been in the hallway on the table charging its battery. An address book was beside it and the burglars had left the hall light switched on.

The coffee maker had been in the kitchen next to the tea, sugar jars and cappuccino mugs.

The china plate had been in the display case in the living room. It was the last remaining piece from a set that had been in the family for years.

Smith's House

The Smith's house had been burgled at Christmas. Mr. and Mrs. Smith and their daughter were out visiting relatives while their son was at a friend's house playing computer games. The Smith's son arrived back home first to discover that they had been broken into. The rest of the family

arrived back only a few minutes later at 10 p.m. The police believe that the burglars broke in through the window in the master bedroom. The police suspect that the burglars used the ladders that were in the back garden in order to reach the window. No one saw the burglars break in although neighbors said that they had seen two individuals in a black van parked in front of the Smith's house at around 8 p.m. that evening. The neighbors had noticed the van when they were going to walk their dog although it had gone by the time they had returned. Attached is a list of items that were stolen (underlined).

The computer had been in the study which is situated at the back of the house. It was on a desk but the burglars had left the keyboard behind.

The video recorder had been on the floor in the sitting room. Movie magazines had been thrown about the room and one of the burglars had dropped a pack of Wrigley's gum.

The telescope had been temporarily set up in the conservatory next to several astronomy books and maps. It was in there while the attic was being renovated.

The rollerblades had been in the son's room. They were lying on the floor with the knee and elbow pads.

The necklace had been in the kitchen. It had been left by the sink in its presentation box. The burglars had smashed several plates on the floor.

The Discman had been on the dresser in the master bedroom. The burglars had knocked the headphones for it on to the floor.

The camcorder had been in the dining room and was lying on the table with the film of a friend's wedding.

The leather coat was in the hallway. It was with the rest of the jackets on the coat rack by the front door. It had belonged to a friend.

The crystal vase had been on the Welsh dresser in the sitting room. The flowers that had been it were thrown over the floor.

The microwave had been in the kitchen. It had been set up on top of the fridge freezer.

Appendix B

Non-critical item questions: Jones' household

1. How did the burglars gain access to the Jones' house?
2. What time of year was it when the Jones' were broken into?
3. How many children do the Jones' have?
4. At what time did the Jones' neighbors see the van?
5. Where were the Jones' neighbors going to when they saw the van?
6. What object did the burglars use to climb up to reach the window of the Jones' house?
7. Who were the Jones' children visiting?
8. At what time did the Jones' get back?
9. How many burglars broke into the Jones' house?
10. Where was the van parked that the Jones' neighbors saw?
11. What color was the van that the Jones's neighbors saw?

Critical item questions: Jones' household

1. The burglars gained access to the Jones' house through the study window. They knocked an item over that was for the scanner. What was this item?
2. When the burglars stole the Jones' briefcase from their hallway, they also left the hallway light switched on but who was the first home to find this out?
3. When the burglars stole the Jones' sculpture from their dining room they also stole the contents of the wallet. Where did they leave the empty wallet?
4. When the burglars stole the alarm clock that was on the bedside table in the Jones' master bedroom, they broke an item. What was broken?

Non-critical item questions: Smith's household

1. How many burglars broke into the Smith's house?
2. Who were the Smith children playing with?
3. What time of year was it when the Smiths were broken into?
4. Where was the van parked that the Smith's neighbors had seen?
5. How many children do the Smith's have?
6. What color was the van that the Smith's neighbors had seen?
7. At what time did the Smith's neighbors see the van?
8. How did the burglars gain access to the Smith's house?
9. What object did the burglars use to climb up to reach the Smith's window?
10. At what time did the Smith's get back?
11. Where were the Smith's neighbors going to when they saw the van?

Critical item questions: Smith's household

1. The burglars gained access to the Smith's house through the master bedroom window, knocking the headphones for the Walkman off the dresser. Where were the Smith's?
2. When the burglars stole the earrings that were next to the sink in the kitchen, they knocked some items on to the floor breaking them. What did they break?
3. When the burglars stole the skateboard from the children's bedroom they knocked some safety equipment off the shelf. What were they?
4. When the burglars stole the DVD player from the living room they threw movie magazines over the room. They also dropped a pack of chewing gum. What was the brand?

Received September 24, 2001
Revision received March 12, 2002
Accepted March 12, 2002 ■