Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Memory and Language 51 (2004) 399–417
Journal of Memory and Language
www.elsevier.com/locate/jml

Transfer appropriate forgetting: The cue-dependent nature of retrieval-induced forgetting☆

Timothy J. Perfect,[a,*] Louisa-Jayne Stark,[a] Jeremy J. Tree,[b] Christopher J.A. Moulin,[c] Lubna Ahmed,[a] and Russell Hutter[d]

[a] *School of Psychology, University of Plymouth, Drake Circus, Plymouth PL4 8AA, UK*
[b] *School of Psychology, University of Exeter, Exeter, EX4 4QG, UK*
[c] *School of Psychology, University of Leeds, Leeds, LS2 9JT, UK*
[d] *School of Psychology, University of Birmingham, Birmingham, B15 2TT, UK*

Received 19 February 2004; revision received 1 June 2004
Available online 21 July 2004

Abstract

Retrieval-induced forgetting is the failure to recall a previously studied word following repeated retrieval of a related item. It has been argued that this is due to retrieval competition between practiced and unpracticed items, which results in inhibition of the non-recalled item, detectable with an independent cue at final test. Three experiments were conducted in which two cues were associated with a target item at encoding. All three studies demonstrated retrieval-induced forgetting when the same retrieval cue was present at practice and test, but not when the second encoding cue was used as an independent probe at final test. These data are not compatible with a general inhibitory account of retrieval-induced forgetting, but support a context-specific account of the phenomenon.

Keywords: Retrieval induced forgetting; Episodic memory; Inhibition; Context

Introduction

In order to have a retrieval system that functions efficiently it is critical not only to recall what is desired, but to not recall what is unwanted. Given the content-addressable nature of human memory, this feat is quite an achievement. Consider, for example, answering the following question: "Where did you go on your last holiday?" Presumably, each of us has experienced a number of holidays that differ in their availability to recall. For example a trip abroad, lasting several weeks may be more memorable than a short break in our own country. Thus when we are asked this question, many potential memories may be activated, and we have to select between them in order to provide an answer to the question.

How does the memory system quickly and efficiently achieve this selection process, and what is the consequence of this? The present work focuses on one putative aspect of the selection process, namely retrieval inhibition. In answer to the question about our most recent holiday, numerous holiday memories might be

☆ The authors thank the Leverhulme Trust (RFG/2/2000/237) and BBSRC (321/S17781) who partially funded the work reported in the manuscript through a grant to the first author. Parts of this work were previously presented at the Experimental Psychology Society Meeting in Bristol, 2000, and the Third International Conference on Memory, Valencia, 2001. The authors thank Colleen Kelley, Michael Anderson, and Reed Hunt for their constructive comments on a previous version of the manuscript.
* Corresponding author. Fax: +44-1752-233362.
E-mail address: tperfect@plymouth.ac.uk (T.J. Perfect).

0749-596X/$ - see front matter © 2004 Elsevier Inc. All rights reserved.
doi:10.1016/j.jml.2004.06.003

potentially available to recall, but only one is the most recent. Thus, it might aid selection of the desired item if the activated memories that do not fit the requirements of the memory search are deactivated, whilst at the same time allowing the activation of the desired memory to accrue. Such mechanisms—of activation and competition between competing representations—are at the heart of many distributed models of human memory and language (e.g., McClelland & Rumelhart, 1986).

The particular focus of the present research is on the long-term consequences of this retrieval competition. It has been repeatedly shown that following retrieval competition, the non-retrieved item becomes less available to recall from long-term memory (Anderson & Bell, 2001; Anderson, Bjork, & Bjork, 1994; Anderson, Green, & McCulloch, 2000; Anderson & McCulloch, 1999; Ciranni & Shimamura, 1999; Koutstaal, Schacter, Johnson, & Galluccio, 1999; MacLeod, 2001; MacLeod & Macae, 2001; Macrae & MacLeod, 1999; Shaw III, Bjork, & Handal, 1995; Smith & Hunt, 2000). Further, it has been proposed that any explanation of this poorer recall performance requires an inhibitory component that acts upon the memory representation itself, rather than through any associative process such as unlearning (Anderson, 2003; Anderson et al., 2000; Anderson & Spellman, 1995; Levy & Anderson, 2002). Evidence for this assertion has come from research using the retrieval-induced forgetting paradigm.

The retrieval-induced forgetting paradigm

The retrieval-induced forgetting paradigm was first introduced by Anderson et al. (1994). The paradigm has three critical phases. In the study phase, participants are asked to learn category cue—exemplar pairs for a later memory test. Critically, the list they see contains several members of each of a number of categories. So, they might see a list that includes the items FRUIT—*orange*, FRUIT—*apple*, TOOL—*hammer*, and TOOL—*spanner* amongst others.

The second phase is the retrieval practice phase, and is presented to the participants as a memory test for the material they have just studied. They are cued to recall some of the items from some of the categories, by means of a category plus word stem cue, such as "FRUIT—or____?" for *orange*. This stage is designed to elicit retrieval for a subset of items, which become designated the practiced (RP+) items. This creates a subset of items that come from the same categories that were not practiced (e.g., *apple*). These items are conventionally designated RP− items. Finally there are yet further items which are not practiced, but which are members of categories which were not practiced (e.g., *hammer*). These items are called either NRp (no retrieval practice) or U (unpracticed) items. We will use the latter term here.

After a 20 min filled delay the final stage of the paradigm is conducted, which is a category cued recall test for all the items that were studied. Anderson et al. (1994) demonstrated that the effect of repeatedly retrieving some members of a category from the category cue was to reduce the likelihood of recalling the unpracticed members of that category. That is, compared to the baseline of recall of U items, RP− items were less likely to be recalled. Thus, retrieving the item *orange* has the counterintuitive effect of making the related item *apple* less likely to be remembered 20 min later.

Anderson et al. (1994) argued that their data were supportive of an inhibitory account of the retrieval-induced forgetting paradigm. Of particular relevance to this argument were the effects on strong (highly available) and weak (less available) members of practiced categories. Across the three experiments reported, strong members of a category showed a 15% recall benefit when they were RP+ items and a 10% impairment when they acted as RP− items, relative to recall of baseline, U items. For weak members of a category, there was a 21% benefit for RP+ items, but no impairment for those items when they were RP− items (in fact there was 3% recall benefit). The presence of retrieval-induced forgetting for strong items occurred whether strong or weak items were practiced. Similarly, the lack of retrieval-induced forgetting for weak items was also independent of the strength of the practiced items. Anderson et al. argued that this pattern is inconsistent with strength dependent competition models, but instead favours an account based upon suppression of the non-practiced item. In this account, stronger category members get inhibited because they are more likely to compete for retrieval, whatever the strength of the practiced item. Conversely, weaker items are unlikely to compete for retrieval during practice, and so there is no need to inhibit them. Anderson et al. (1994) concluded that their results "..are consistent… with a process of active suppression, applied directly to competing items to the extent that those items interfere with task demands." (p. 1080).

Cue independence as evidence for an inhibitory account

Whilst Anderson et al. (1994) argued for an inhibitory account of their findings, Anderson and Spellman (1995) acknowledged that there are several non-inhibitory accounts that could potentially explain those data. (For reviews of these non-inhibitory accounts see Anderson & Bjork, 1994; Anderson & Neely, 1996). For example, it is possible that the repeated retrieval practice of FRUIT—*orange* is likely to strengthen the association between the cue FRUIT and the exemplar *orange*, such that when FRUIT reappears as a cue, then *orange* is the most likely item to be recalled. Indeed, it is possible that *orange* is so strongly available that other items can no longer be recalled; in this model access to *apple* may be occluded by the repeated access of the item *orange*. Notice that the failure to recall *apple* in this model says nothing about the strength of the memory representation of *apple*, nor about the strength of the association between FRUIT and *apple*. The observed forgetting is caused merely by the strengthening of the FRUIT—*orange* link, together with a retrieval process that is resource limited. That is, only one item can be recalled at a time, and retrieval attempts will cease if the desired answer is not found quickly enough.

A different form of associative model based upon the principle of unlearning (Melton & Irwin, 1940) might also suffice. In such a model, the repeated retrieval of *orange* leads not only to the strengthening of the FRUIT—*orange* association, but also to the weakening of the association between FRUIT and *apple*. Thus when the cue FRUIT appears, *apple* is less likely to be cued, and so is forgotten. Note here, as before, the representation *apple* is not inhibited, since it is the FRUIT—*apple* link that is weakened, not the representation of *apple* itself.

Anderson and Spellman (1995) argued that the only way to distinguish between inhibitory and associative models of retrieval-induced forgetting is to use an independent retrieval cue at the final test stage (see also Levy & Anderson, 2002). They argued that all the non-inhibitory accounts share an assumption that the forgetting is due to a failure of the cue used at practice to elicit the target at test. Thus, the forgetting effect should be seen only if the same cue is used at the final test phase as was used in the practice phase. However, if retrieval-induced forgetting occurs because the non-retrieved item is inhibited at practice, then it follows that the forgetting should be seen with any cue, since it is the memory itself that is inhibited, rather than the link with a specific practice cue. This is the basis of the independent probe version of the retrieval-induced forgetting paradigm introduced by Anderson and Spellman (1995), and extended by Anderson et al. (2000).

In the independent probe technique, the essential elements of the retrieval-induced forgetting paradigm are retained, with the only modification used being the final test for the putatively inhibited item, which is cued by a different word to that used to elicit the practiced item. Two different forms of independent probe techniques have been used. In Anderson and Spellman (1995), the effects of retrieval practice of an item in one category (e.g., GREEN—*emerald*) were examined on retrieval of an item paired with a different category cue at encoding (e.g., SOUP—*mushroom*). In Anderson et al. (2000), the effects of practice of an item from a category (e.g., RED—*brick*) were tested for other items from the same category (e.g. *tomato*), but with the independent category cue FOOD used at the final test stage. We discuss each methodology in turn.

In Anderson and Spellman's (1995) paradigm, participants studied lists of category-exemplar pairs in the manner introduced by Anderson et al. (1994). However, within the items studied there were systematic relations between the items. So, in Experiment 2, for example, participants studied 6 items associated with the cue GREEN and 6 with SOUP. Critically, half of the GREEN items were vegetables, as were half of the SOUP items. Thus, there was an implicit category of VEGETABLES within the experiment, although the cue VEGETABLES itself was never presented. In the retrieval practice phase, participants practiced retrieval for the non-vegetable items from one of the presented categories (e.g., GREEN—*emerald*). This results in four kinds of items at test. As in Anderson et al. (1994) there were practiced items (RP+), and unpracticed items from the same category (RP−). However, whereas Anderson et al. (1994) merely had unpracticed items from unpracticed categories (NRp items), Anderson and Spellman's (1995) methodology led to NRp items that are similar to the RP− items (designated NRpS), and NRp items that are dissimilar to the RP− items (NRpD items).

The results indicated that retrieval practice of GREEN—*emerald* led to retrieval-induced forgetting for the other items associated with that category, which is the standard retrieval-induced forgetting effect seen for the RP− items (e.g., GREEN—*lettuce*). In addition, forgetting was also seen for the items from the other category that were similar to the RP− items (i.e., NRpS items such as *mushroom*), even when cued with SOUP. Thus, since forgetting was seen for the NRpS item using a non-practiced category cue, Anderson and Spellman (1995) argued that the forgetting effect must be due to inhibitory processes acting upon the memory representations, rather than due to an associative process that links the category and the item. They proposed a feature competition explanation of this pattern, in which the cue (GREEN) elicits the desired targets (e.g., *emerald*), whilst at the same time inhibiting those features of the undesired items (e.g., *lettuce*) that are not shared with the practiced items. All three non-retrieved items are vegetables, and so the vegetable qualities of these unwanted green items are inhibited. These inhibited vegetable-features will be shared with other vegetables, such as mushroom, and so the result is that *mushroom* is also inhibited. Thus *mushroom* is difficult to recall, even when cued with SOUP.

Whilst this experimental methodology is elegant, and the argument persuasive, we reserve judgment about the inhibitory account of retrieval-induced forgetting observed in this study. One reason is that the results are perhaps surprising given the previous findings reported by Anderson et al. (1994). They demonstrated that weak category exemplars (e.g., FRUIT—*nectarine*) were not inhibited by practice of other category members, whether they are strong (e.g., *orange*), or weak (e.g., *fig*) members of the same category. Thus according to Anderson et al. (1994) practice of *orange* does not inhibit concepts as similar as *nectarine* or *guava*. However, in Anderson and Spellman (1995) practice of *dollar*, *emerald* and *lawn* from the cue GREEN produces forgetting for *artichoke*, *lettuce*, and *pepper* from the same cue, which in turn produces forgetting of *mushroom*, *onion*, and *tomato* from the cue SOUP. Although norms are not available, it does not appear likely that *artichoke*, *lettuce*, and *pepper* are strong members of the category GREEN. Thus, following Anderson et al. (1994) they would not be expected to compete for retrieval from the cue GREEN, and so should not be inhibited. Thus, in the context of the Anderson et al. (1994) findings, the inhibition reported in Anderson and Spellman (1995) appears unexpected given the materials used.

A second concern arises from reconsideration of Anderson and Spellman's (1995) data across three crucial experiments. Experiments 2 and 4 in their paper demonstrated second-order inhibitory effects along the lines discussed above. In each case, the comparison was made between conditions in which the critical NRpS (*mushroom)* item was studied either in the context of related or unrelated RP− items. In each case, recall was lower in the presence of similar RP− items, thus supporting Anderson and Spellman's claims of inhibition being spread because of similarity between RP− and NRpS items. The key difference between the two experiments was that Experiment 4 controlled for output interference effects, which the authors acknowledge was a potential confounding explanation for the data reported for Experiment 2.

Table 1 reproduces the data from Anderson and Spellman (1995), taken from Tables 2 and 4 of their paper. As well as providing the critical comparisons that Anderson and Spellman focused on, this table also contains data from Experiment 3a, which was conducted without the key retrieval practice phase. This experiment enabled the authors to demonstrate that the previous forgetting effects seen in related conditions were not because of some aspect of list structure across related and unrelated lists. However, this condition also provides a useful baseline against which the previous effects on Rp− and NRpS items can be interpreted, since in that experiment there was no retrieval practice, and so no retrieval-induced forgetting should occur. This baseline condition produced a recall rate of 48% for both related and unrelated conditions.

The crucial data concern the recall rates for NRpS items, since they were used by Anderson and Spellman (1995) to argue for an inhibitory account of retrieval-induced forgetting. These data are boxed in Table 1, although performance on the other critical items is included for completeness. First, consider the data from the related conditions, in the upper box, which are the conditions in which retrieval-induced forgetting is predicted. Compared to the baseline recall rate, retrieval practice produced an apparent forgetting rate of 11% in Experiment 2, where output interference was not controlled, but only 3% where output interference was controlled in Experiment 4. Now, turn to the unrelated conditions, where no retrieval-induced forgetting is predicted, and which were used by Anderson and Spellman as the contrast conditions in Experiments 2 and 4. Since no retrieval-induced forgetting is predicted, one might expect similar performance to the baseline level of 48% observed when no retrieval practice occurred. However, recall of NRpS items was 4% above baseline for Experiment 2, and 9% above baseline in Experiment 4. Thus, the robust retrieval-induced forgetting effects reported by Anderson and Spellman for Experiments 2 and 4, obtained by contrasting the related and unrelated conditions in each experiment, seem to be due to an unexpectedly high level of recall for the NRpS

Table 1
Mean percentage recall in Anderson and Spellman (1995) Experiments 2 and 4, for related and unrelated conditions, with or without control for output interference, compared to baseline (no retrieval practice) performance in Experiment 3a

	Item status				
	Rp+ *Emerald*	Rp− R *Lettuce*	Rp− U *grenade*	NRpS *Mushroom*	NRpD *chicken*
Related conditions					
Experiment 2	65	36		37	43
Experiment 4	67		41	45	44
Experiment 3a				48	44
Unrelated conditions					
Experiment 2	72		42	52	43
Experiment 4	69	43		57	47
Experiment 3a				48	36

Notes. Experiments 2 and 4 used a second-order inhibition paradigm. However, only Experiment 4 controlled for output interference by fixing retrieval order. The data above are those reported when unpracticed items are reported first, and so output interference is minimized. Experiment 3 involved recall of the same items as Experiments 2 and 4, but with retrieval practice of filler items rather than critical items.

items in the unrelated condition, rather than due to suppression in the related condition. Whilst it is beyond the scope of the present discussion to fully explain this pattern, it is clear that the majority of the effect reported by Anderson and Spellman (1995) is not consistent with an inhibitory effect. Rather, it appears to be consistent with an inexplicable retrieval benefit afforded to NRpS items in the unrelated condition. Unfortunately, the other studies that have demonstrated second order inhibitory effects (Anderson & Bell, 2001) failed to include a no-practice control, and so it is impossible to know whether or not a similar effect could explain those data.

A final concern with this second-order inhibitory effect is that there is one published failure to replicate this effect, despite using a design with greater numbers of items per category, and more participants in the study than in the original Anderson and Spellman (1995) paper (Williams & Zacks, 2001). Whilst this failure to replicate is only an isolated finding, together with the concerns raised above, it seems reasonable to maintain a reasonable level of doubt about the reliability of the second-order inhibitory effect. Since this effect is crucial in establishing inhibition as the mechanism underlying retrieval-induced forgetting, it is clear that further evidence is required.

Although the Anderson et al. (2000) paradigm introduces an entirely new independent probe at the final recall test (i.e., one not presented during study), similar problems may complicate interpretation of these findings. In this paradigm, participants study and practice in the standard manner. However, in the final test condition using an independent probe, participants are tested using semantically related cues that were not present at encoding or practice. So, for example, RED—*brick* and RED—*tomato* are studied. Following retrieval practice of RED—*brick*, *tomato* is less likely to be recalled, either from the cue RED, or the cue FOOD, which was never present at study. In this manner, cue-independent forgetting is demonstrated, and hence inhibition. However, as in Anderson and Spellman (1995), the category RED was studied with multiple (eight) items, half of which were foods. No food items appeared in any other category. Thus food is a highly salient aspect of the RED category, and participants are likely to have noticed this, and to have formed an association between RED and FOOD. Now imagine a participant faced with recalling items from the cue FOOD. If their retrieval strategy is to remember that foods were uniquely associated with RED, then they may use RED as a proxy retrieval cue rather than using FOOD directly. Why should they do this? One reason is that this independent probe technique is effectively a version of an encoding specificity experiment (Thomson & Tulving, 1970; Tulving & Thomson, 1973) in that participants are given a strong associate of an item as a memory cue for an item studied with another cue. However, as Tulving and Thomson (1973) showed, this is a poor cue, relative to the originally encoded cue. Thus, one might expect the cue FOOD not to be a very useful retrieval cue to participants. Given that, what can participants do, other than to try to use a more effective cue? The most recent relevant aspect of the experiment will have been the retrieval practice phase, which means that the category cue RED is highly accessible. Given that this was the cue originally paired with the forgotten item, and that half the studied red things were food, it is possible that at least some participants try to use the cue RED when presented with the cue FOOD. The consequence is that this merely reinstates the originally practiced retrieval cue, and so the non-practiced items are likely to be poorly recalled. Thus an associative account of the supposedly independently cued forgetting effect is possible, based on covert-cueing.

Anderson and his colleagues (Anderson & Bell, 2001; Anderson et al., 2000; Anderson & Spellman, 1995) have argued that such covert-cueing effects cannot explain their data for two reasons. One is that those who report using such a strategy show numerically, but not statistically, less inhibition (Anderson, Bjork, & Bjork, 2000). However, this argument rests upon self-reported use of a covert-cueing strategy in which participants report on a rating scale how often they mentally scanned through previously encountered category names to find the sought after item. However, this can only be completed after recall, and so it is entirely plausible that the reported strategy is driven by the success or failure at recall, rather than vice-versa. That is, those who perform poorly (i.e., show forgetting effects) deny the use of the strategy, whilst those who do well (i.e., show less forgetting) claim that they did use the strategy. Thus, there is good reason to question the usefulness of the self-report data. The second argument is that covert-cueing predicts forgetting where none is found (Anderson & Bell, 2001). For example, participants faced with the non-practiced category SOUP may recall the overlap with the previous category GREEN, via the shared vegetable items. This would then lead them to recall the practiced GREEN exemplars, so blocking recall of both SOUP exemplars. Crucially, this argument applies equally to items that share features (e.g., *mushroom*) as well as those that do not (e.g., *turkey*). Thus forgetting for NRpD items is predicted, but was not observed in either Anderson and Spellman (1995) or Anderson and Bell (2001). However, as we noted above, Anderson and Spellman's (1995) data are not as clear cut as they might be, when no-practice baseline is taken into account, and Anderson and Bell (2001) did not report any no-practice baseline data.

Thus, whilst it is not clear that a covert-cueing account can fully explain the pattern of data in the literature, it is also the case that the inhibitory account is open to challenge. Both independent probe methodologies can be challenged with regards how well they provide an independent cue for the putatively inhibited item. Their use of the term independent refers to the non-occurrence of the critical cue during the first two stages of the experimental paradigm, which thereby produces a cue that is independent of any associations created in the study or practice phases. However, there is a stronger sense of the word independence, which is logical or statistical independence. Two events that are logically independent are completely unrelated, in the sense that the outcome of one event does not influence the outcome of the other. On this criterion the previous methodology fails, since it is clear that the independent probes are not unrelated to the previously experienced events. For example, someone given the cue word VEGETABLE is more likely to go on to generate the item *onion* than they are *siren*, whatever their memory for those two items. Similarly, whatever their memory they are more likely to think of the previous category cue GREEN than the previous category cue LOUD. This dependence between the independent cues and the list structure is an inevitable consequence of relying upon prior semantic associations between independent cues and to-be-remembered items, and associations between the items themselves.

In three experiments presented here, we adopt a different methodology in which to-be-remembered items are associated with an item that is independent in all respects, save for their prior episodic co-occurrence. In Experiments 1 and 2, participants are presented with category—exemplar pairs as in the standard retrieval-induced forgetting paradigm. However, in addition, each pair is presented along with an independent item, namely a photograph of a face. Each face is arbitrarily paired with a category—exemplar pair, and so is formally independent of the to-be-remembered information. That is, there is no information in the face that will cue the particular exemplar beyond the episodic relationship established by their co-occurrence. There is no hidden structure to the categories. Thus, these cues are statistically and semantically independent of the both categorical cues, and all target items. In Experiment 1, the retrieval practice phase replicates that used by Anderson et al. (1994), with retrieval of category exemplars cued by categories and word stems alone. No faces are presented at practice. This means that at the final test stage it is possible to cue memory for the items either by category cue (as in the standard retrieval-induced forgetting paradigm), by the face originally paired with the exemplar, or by a joint category plus face cue. In Experiment 3, we used words with no prior association with either the exemplars or the category cues as independent episodic cues in the same manner. Additionally we removed any association between the episodic cue and the category cue by presenting the episodic cue together with the exemplar words in a separate list, so that there was never any co-occurrence of the episodic cues and the category cues.

The predictions are relatively straightforward for the inhibitory account in each case. Since the retrieval practice phase involves category cueing, competition will be invoked and the non-retrieved items will be inhibited. Consequently, poorer recall should be demonstrated for those RP− items, whatever the retrieval cue, even if that previous cue is one that was only ever episodically associated with the forgotten item. On the other hand, if retrieval-induced forgetting is due to an associative bias caused by retrieval practice, one would expect a different pattern. As before, the use of category cues at test should produce retrieval-induced forgetting, but now face cues should not demonstrate retrieval-induced forgetting effects. The prediction regarding combined cues is less clear cut. If participants rely heavily on the category cues, ignoring the faces at test, then retrieval-induced forgetting may occur. On the other hand, the extent to which the faces are useful cues should be revealed by the reduction in the retrieval-induced forgetting effect relative to the category cue condition. That is, participants may suffer a retrieval-induced forgetting effect due to the categories, but be able to overcome it by use of the faces.

Experiment 1

Method

Participants

A total of 90 participants were tested in this study. All were undergraduate students recruited via posters distributed around the University of Plymouth. The 90 participants were each randomly assigned to one of the three experimental conditions. They participated for either payment, or partial course credit.

Materials

There were 24 category—exemplar word pairs used in the study phase, comprising of six different categories, with four exemplars for each. No two items began with the same two letters. Twenty-four photographs were used, half of which were male, and half female, and they varied in age and race in order to increase distinctiveness. For the verbal materials for this and subsequent experiments, see Appendix.

Procedure

The experiment consisted of the standard phases used in retrieval-induced forgetting studies; a study phase, a retrieval practice phase and a final cued-recall test phase. The study phase and the practice phase were identical for all participants. Each participant was tested individually in a quiet test laboratory. Participants were informed that they would be participating in a memory experiment involving a series of photographs of faces, above each of which there would be a category-exemplar word pair. They were instructed to try and relate each item to its category and to the face, and to remember as many of the words as they could. An example of the type of test item that they were about to see was given, from a category that was not part of the task. Once the experimenter was satisfied that the participant understood the task, the experiment proper was started. The materials were presented on a lap-top computer placed on the desk at a comfortable reading distance for the participant. PowerPoint slide show was used to present the materials. The category cues were shown in upper case and the exemplars were shown in lower case (e.g., SPORT- tennis), in 36 point Arial font. For the study phase, the 24 stimuli (category + exemplar + face) were presented at a rate of 4 s per item, with an inter-stimulus interval of 1 s. At the end of the list, the participants were thanked, and the instructions for the practice phase appeared.

The retrieval practice phase for all participants consisted of retrieval practice of two items from four categories. Retrieval was cued using a category cue plus the initial two letters of the exemplar, as in previous studies (e.g., SPORT—te____). The participants were instructed to complete each stem with the appropriate word from the study list. They were given 4 s per cue to provide a verbal response. The list was repeated three times in random order.

Before the final testing phase the participants completed visual puzzles, taken from popular puzzle books, for a period of 7 min. They were then randomly assigned to one of three testing conditions, which are described separately. In all conditions, test cues were presented in random order on the lap-top computer, and participants wrote down their responses on a response sheet that was provided.

In the *category cue* condition, participants were presented with the category cues without stems. They were instructed to recall as many items as they could remember that were paired with that category cue during the original study phase. Participants were given a maximum of 20 s to write down their responses for each category on the response sheet provided. At the end of the 20 s interval the screen went blank for 1 s before the next category cue appeared.

In the *face cue* condition, participants were presented with each photograph that had appeared during the encoding phase, and asked to generate the exemplar associated with that person. Participants were asked to write down their responses on the response sheet provided, and had a maximum of 5 s per item. The 24 face cues were shown in a randomized order with a 1s inter-stimulus interval between each cue.

In the *joint face plus category cue* condition, participants were presented with both the face and the category cue originally presented with the exemplar, and asked to recall the item that was associated with that person. Once again they were given a maximum of 5 s per item in which to write down their response on the response sheet provided. The 24 face plus category cues were shown in a randomized order with a 1 s inter-stimulus interval between each cue.

Results

The strategy for data analysis throughout was to examine the interaction between item status (RP+, RP−, and U) and condition (category cue, face cue, and joint cue) using a two-way ANOVA. Because the focus was on retrieval inhibition, any significant interaction was followed up by a comparison of recall rates in the RP− and U conditions, separately for each of the three conditions, to enable comparison with previous experiments. A one-tailed α level of .05 was used for these comparisons. However, before we present these data, it is necessary to explain the two scoring criteria that were used. In the strict scoring procedure used with the face and joint face plus category cues, responses were only coded as correct if the exemplar that was recalled had been paired with the particular face shown on a study trial. A concrete example makes this point clearer. Imagine that during study, Face A appeared with *SPORT—tennis*, Face B with *SPORT—rugby*, and Face C with *FOOD—cheese*. At recall, when presented with Face A, only tennis is the correct answer. However, this makes comparison with the category cued condition problematic, since the cue SPORT does not specify which particular item (*rugby* or *tennis*) must be recalled first. For this reason, a looser scoring criterion was also used in which recall of any appropriate category member to the face was rewarded. Under this criterion, when presented with Face A, recall of either *tennis* or *rugby* was counted as correct, but recall of *cheese* was not. Of course, if participants recalled the same item to different faces (*rugby* to both Face A and Face B) they were only credited with recall of a single item.

The interpretation of the two scoring methods is relatively straightforward with regards the inhibitory account of retrieval-induced forgetting. If an RP− item is inhibited, it should not be recalled to any retrieval cue, even one with which it was not initially paired. Or, put another way, if participants are recalling an item from the wrong cue, it is hard to argue that the item is inhibited. On the other hand, large differences in retrieval-induced forgetting observed between the two scoring methods would be indicative of an associative process, since it would suggest that items not recallable from one cue are available to another cue.

Table 2 shows the mean recall performance for each kind of item for each retrieval cue, as scored by the strict recall procedure. Overall, there were main effects of item status, $F(2, 174) = 35.8$, $MSE = 82.1$, $p < .001$, condition $F(2, 87) = 6.17$, $MSE = 34.4$, $p < .01$ and critically, a significant interaction between the two, $F(4, 174) = 13.7$, $MSE = 31.5$, $p < .001$. Both main effects and the interaction term remained significant when scored using the looser condition also (all Fs > 9). The magnitudes of the retrieval-induced forgetting effects were therefore followed up separately for each condition. Where appropriate, we report the results of the looser scoring criterion in the text.

The category cue condition: There was a robust retrieval-induced forgetting effect, $t(29) = 2.60$ in this condition. Therefore this replicates the standard retrieval-induced forgetting effect with category cues, notwithstanding the presence of the faces at encoding.

The face cue condition: There was no reliable retrieval-induced forgetting effect, $t(29) = 1.2$. Using the looser criterion, participants recalled an average of 3.90 (*SD* 2.34) RP+ items, 3.53 (*SD* 2.11) RP− items and 3.73 (*SD* 1.78) U items. As with the strict scoring criterion, there was no reliable forgetting effect using faces, $t(29) = 0.57$.

Joint face and category cue condition: There was no retrieval-induced forgetting effect, $t(29) = 0.07$. Use of the loose scoring criterion forgetting had no impact on the observed retrieval-induced forgetting effect, $t(29) = 0.40$. For RP+ items, recall now averaged 6.43 (*SD* 1.50). The equivalent figures for RP− and U were 4.43 (*SD* 1.96) and 4.60 (*SD* 1.63).

The data from this experiment were clear cut, and appear to contradict the claim that retrieval-induced forgetting is cue independent. When precisely the same retrieval cue was used in the practice and test phases, robust retrieval-induced forgetting was observed. When different retrieval cues were used, no reliable retrieval-induced forgetting was observed. The most telling comparison comes from comparing the category cue condition and the joint face plus cue condition. These two conditions differ only in the presence of the face at test. When the category cue alone was presented, participants recalled an average of 3.03 (37.9%) RP− items, compared to 4.17 (52.1%) U items under the strict scoring criterion. The equivalent figures for the joint face and category cues were 4.0 (50.0%) and 4.03 (50.4%). The additional face cue therefore had no benefit to recall of the U items but increased recall of the RP− items by 12.1%. Exactly the same pattern is obtained if the loose scoring criterion is used. It is hard to escape the conclusion that the RP− items are only unavailable to recall from the category cue. Providing the face cue at test makes available RP− items that would not have been recalled from the category cue alone. The retrieval-induced forgetting effect appears to be entirely cue-dependent.

One complication in interpretation of these data is the fact that performance on the U items varies across experimental conditions. Whilst this is understandable, it does allow the possibility that the forgetting effects are mediated by the overall level of performance in the baseline (U) condition. In order to explore this possibility, we conducted a post hoc analysis of the performance of the three groups, by selecting sub-samples that were matched on their performance in the baseline condition. To do this, we dropped the five worst performing people in the face-cue condition, and the three best performing people in the category cue condition. As Table 3 shows, dropping just eight participants from 90 produced three groups who were matched on performance for baseline items. However, performance was not matched for RP− items, $F(2, 79) = 4.32$, $MSE = 3.40$, $p < .017$, with post hoc tests indicating that RP− performance in the category cue condition was marginally worse than in the face cue condition ($p < .052$) and significantly worse than in the joint-cue

Table 2
Mean (and *SD*) number of items recalled for practiced items from practiced categories (RP+), unpracticed items from practiced categories (RP−), and unpracticed items from unpracticed categories (U), dependent upon retrieval cue used, for Experiments 1 and 2

Retrieval cues	Item status					
	RP+		RP−		U	
	Mean	*SD*	Mean	*SD*	Mean	*SD*
Experiment 1: Category cues at practice						
Category	7.10	0.80	3.03	1.86	4.17	1.60
Face	3.63	2.37	3.20	2.25	3.63	1.83
Joint	5.07	2.20	4.00	1.91	4.03	1.81
Experiment 2: Joint face plus category cues at practice						
Category	7.47	0.63	3.37	1.56	4.63	1.79
Face	6.37	1.71	3.50	2.37	3.17	2.29
Joint	7.00	1.14	4.03	2.39	4.57	1.94

Notes. Category, category cue used at retrieval. Face, face cue used at retrieval. Joint, joint face plus category cue used at retrieval.

Table 3
Mean (and *SD*) number of items recalled for practiced items from practiced categories (RP+), unpracticed items from practiced categories (RP−), and unpracticed items from unpracticed categories (U), dependent upon retrieval cue used, for Experiments 1 and 2, for subsamples matched on performance on U items

Retrieval cues	Item status					
	RP+		RP−		U	
	Mean	*SD*	Mean	*SD*	Mean	*SD*
Experiment 1: Category cues at practice						
Category (27)	7.11	0.80	2.59	1.34	4.07	1.57
Face (25)	4.24	2.10	3.60	2.19	3.96	2.20
Joint (30)	5.07	2.20	4.00	1.91	4.03	1.81
Experiment 2: Joint face plus category cues at practice						
Category (29)	7.45	0.63	3.27	1.59	4.55	1.76
Face (17)	7.41	0.71	5.06	1.74	4.58	1.97
Joint (30)	7.00	1.14	4.03	2.39	4.57	1.94

Notes. Category, category cue used at retrieval. Face, face cue used at retrieval. Joint, joint face plus category cue used at retrieval. Numbers in parentheses represent the number of participants in each condition, matched on performance on U items.

condition ($p < .005$). This analysis therefore confirms the pattern seen earlier. In a sample matched on baseline performance across conditions, the degree of forgetting observed across conditions differs significantly, with reliable forgetting observed only with category cues, $t(27) = 3.66$, $p < .001$, and absent for face cues, or joint cues ($t < 1$ in each case).

One notable feature of the performance in the face-cued condition was that no retrieval practice benefit was observed. That is, RP+ recall performance does not exceed baseline, U item recall. In this context, perhaps the absence of retrieval-induced forgetting might be considered less surprising, and so the lack of a retrieval-induced forgetting effect with the face cues is less compelling. This observation raises several issues for consideration. First, and most obvious is the fact that the joint face plus category cue condition did show a retrieval practice effect, yet showed no forgetting, which contradicts the simple notion that the lack of forgetting is simply due to a lack of practice effect (see also Anderson et al., 2000). One must also ask what the lack of a retrieval practice benefit means, given the robust demonstration of such benefits when category cues are used at the final test. Given that the conditions only differ in the cues used in the final test, why is no retrieval practice benefit seen when faces are used as cues? One answer is that, in line with Tulving and Thomson's (1973) encoding specificity principle, the amount of recall seen is a consequence of the overlap between encoding and test. Given that the category-cued practice phase overlaps little with the final face-cued recall test, the lack of benefit of practice is perhaps less surprising, and consequently the same may apply to the lack of forgetting effect. We return to this issue in the General Discussion.

Notwithstanding these comments, it might be more persuasive to demonstrate the lack of retrieval-induced forgetting to face cues in conditions that show retrieval practice benefits to those cues. Thus, in the second experiment we decided to include the face cues during the retrieval practice phase, in order to increase the salience of the faces as retrieval cues. We anticipated that this would produce a retrieval practice effect observable with face cues at test. However, the consequences for the forgetting effect are less clear cut.

What does the inhibitory account of retrieval-induced forgetting predict from such a procedure? A prerequisite for retrieval-induced forgetting is prior retrieval competition. Since each face is uniquely associated with an exemplar, the presentation of the faces with the categories during retrieval practice should reduce retrieval-induced forgetting seen in this study, relative to Experiment 1, perhaps removing it entirely. On the other hand, it might reasonably be expected that, given the pattern of performance for baseline U items in Experiment 1, faces are weaker cues than category labels. Hence, during retrieval practice, participants may rely on the category plus word stem cues for retrieval, and hence induce competition, and consequently retrieval-induced forgetting. From an inhibitory standpoint, this competition should result in poor recall for the RP− items from any retrieval cue. Thus, according to an inhibitory account of retrieval-induced forgetting, if any forgetting is seen, it should be seen equally across all test conditions.

Experiment 2

Participants

A total of 90 participants were tested in this study. All the participants were either undergraduate or postgraduate students recruited via posters distributed around the University of Plymouth. None of them had participated in Experiment 1. Participants were randomly assigned to one of three experimental conditions. No other biographical data were recorded.

Procedure

The methodology for this study was identical to the previous study, except for the practice phase. During practice participants practiced retrieval of the RP+ items from the combined face plus category cue used at encoding.

Results and discussion

The data were scored and analyzed as in the previous experiment. A 3 (item status: RP+, RP, and U) $\times$ 3 (condition: face cue, category cue, joint cue) ANOVA revealed main effects of item status, $F(2, 174) = 153.5$, $MSE = 287.5$, $p < .001$, and condition, $F(2, 87) = 3.24$, $MSE = 20.9$, $p < .05$, and a significant interaction between the two, $F(4, 174) = 3.36$, $MSE = 6.29$, $p < .05$. The analyses based on loose scoring produced the same pattern (all Fs > 5.9). As in Experiment 1, the follow up analyses focused on the retrieval-induced forgetting effect (RP− vs U) in each condition. Mean performance across condition, scored by strict scoring method, is shown in the bottom half of Table 2, and where appropriate results from the loose scoring criterion are discussed in the text.

The category cue condition: As the mean recall performance shown in Table 2 indicates, there was a reliable retrieval-induced forgetting effect, $t(29) = 3.28$. Therefore, as in Experiment 1 we demonstrated a reliable retrieval-induced forgetting effect using the standard category cues at test. This occurred notwithstanding the presence of the faces at study and practice.

The face cue condition: Using the strict scoring criterion resulted in numerically higher recall performance for RP− items than for U items, and so there was no reliable retrieval-induced forgetting effect. The loose scoring criterion showed the same pattern, with an average of 6.43 (*SD* 1.71) RP+ items, 3.73 (*SD* 2.33) RP− items, and 3.36 (*SD* 2.17) U items being recalled. Thus, as in Experiment 1, cueing with a face showed no evidence of a retrieval inhibition effect. However, unlike the previous study, there was strong evidence of a practice effect in this experiment, and hence the absence of a retrieval-induced forgetting effect with face cues cannot be attributed to lack of a practice effect.

Joint face plus category cue condition: Using the strict recall criterion revealed a marginally significant retrieval-induced forgetting effect by one tailed test, $t(29) = 1.56$, $p < .07$. Using the looser scoring criterion produced average recall of 7.33 (*SD* 0.92) RP+ items, 4.73 (*SD* 2.06) RP− items and 5.80 (*SD* 1.69) U items which constitutes a significant retrieval-induced forgetting effect, $t(29) = 3.06$.

There are a number of noteworthy points about these data. First, we replicated the retrieval-induced forgetting effect when category cues were used at test. Whilst the category cued test is the standard procedure, and has been used to demonstrate retrieval-induced forgetting in previous studies, the presence of the effect was by no means certain here. This follows because the forgetting effect is posited to be due to competition at the retrieval practice phase. However, in this study, there was a second retrieval cue available at practice, which could have served to reduce competition. That is, the face alone unambiguously specified the required target, and so competitor targets might not have been accessed, and so might not have had to be inhibited. However, the presence of the forgetting effect suggests otherwise and suggests that the face cue was relatively weak compared to the category cue. Thus it appears that participants were largely using the category cue as a retrieval cue in the practice phase, which had the ironic effect of subsequently producing a forgetting effect for the related items.

We also replicated the previously observed lack of retrieval-induced forgetting when faces were used as the final retrieval cues. However, unlike the previous study, here the face cues produced better performance for the RP+ items, as would be expected given the role of faces in the retrieval practice phase. The absence of a retrieval inhibition effect is once again problematic for an inhibitory account of retrieval-induced forgetting, given performance in the category-cued condition.

However, perhaps the most interesting aspect of the second study was the reliable retrieval-induced forgetting effect observed with the joint cues. In the previous study, when no faces had been used at test, no forgetting effect was seen in the final test when joint cues were used. However, here, when the joint cues were used at practice, joint cues at test did produce the effect. Clearly, the presence of a forgetting effect is compatible with the inhibitory account of forgetting, but it is hard to square this account with the lack of such an effect in the previous study. Why should we obtain a forgetting effect when joint retrieval cues were used at test in Experiment 2, but not in Experiment 1?

The basis of the inhibitory account is that the retrieval cues at practice induce competition, and that this competition produces inhibition that will result in forgetting to any cue. However, two elements of the pattern with joint cues contradict this theoretical account. First, the presence of the faces at practice reduced the potential for competition in Experiment 2, relative to Experiment 1, yet a forgetting effect was only seen for joint cues in the second experiment. Second, if the forgetting effect is cue independent then it should be seen with any retrieval cue, yet identical test cues result in forgetting in Experiment 2, but not in Experiment 1.

As in the previous study, comparison of the category cues with the face plus category cues is particularly informative. Under the strict scoring procedure, category cues elicited 3.37 (42.1%) RP− items, and 4.63 (57.9%) U items. Relative to this, the joint cues showed recall levels of 4.03 (50.4%) RP− items and 4.57 (57.1%) for the U items. Thus, under strict scoring, the addition of the faces to the category cues only benefits recall of RP− items, and does not benefit recall of U items at all. This pattern resembles the previous study. However, the pattern of findings is slightly different with the loose scoring criterion. Here, the joint cues elicited 58.8% recall of RP− items, and 72.5% of U items. Thus, the addition of the face cues to the category cues benefits both the RP− and U items equally. This suggests that the items that cannot be recalled to the category cues remain inaccessible to the joint category plus face cues, which is in contrast to the previous experiment.

As in Experiment 1, baseline differences in performance may complicate interpretation of these data. Consequently, we replicated our previous strategy of selecting sub-samples of each group, matched on performance for the U items. This resulted in the selection of the best 17 participants in the face cue condition and the best 29 in the category condition to be compared to the full sample in the joint cue condition. Their performance is shown in Table 3. As in Experiment 1, in this sample matched on performance for U items, there was a significant difference in performance for RP− items, $F(2,73) = 4.50$, $\text{MSE} = 3.80$, $p < .014$. Post hoc testing indicated that recall using face cues was significantly better than recall using category cues ($p < .004$) and marginally better than recall using joint cues ($p < .09$), but there was no difference between category cues and joint cues ($p < .14$). Whilst the category cue sub-sample show reliable retrieval-induced forgetting, $t(29) = 3.19$, $p < .003$, the face cue condition showed numerically superior performance for RP− items over U items, and so showed no forgetting effect. Thus, the result of the analyses of baseline-matched sub-samples supports the conclusions drawn from the full sample: the degree of forgetting observed is a function of the retrieval cues used at test.

These data can be explained in terms of transfer appropriate forgetting. That is, forgetting is seen when the retrieval conditions most closely match the conditions of the first retrieval competition. Hence, in Experiment 1, the retrieval competition was induced by category cues alone, and only category cues elicit the forgetting effect. In Experiment 2, where the retrieval practice phase involved both category cues and faces, then the same combination of cues produce the forgetting effect. In neither case did the faces alone produce the forgetting effect. We will expand on this notion further in the General Discussion that follows the final experiment.

One potential criticism of the first two experiments is that output order was not controlled sufficiently in the final tests of memory, and thus there is potential for output interference to explain our findings (cf. Anderson et al., 1994; Anderson & Spellman, 1995). Output interference is a potential confound because stronger members of a category are likely to be output first, and during retrieval blocking of subsequent items can occur. This is particularly pertinent to the retrieval-induced forgetting paradigm since the retrieval practice phase will lead to strengthening of the RP+ items, and so in free recall these will be output first, producing greater output interference for the RP− items than will occur, on average, for the U items. This would explain why retrieval-induced forgetting is seen in the category cued recall performance in both experiments.

However, there are aspects of the data that are incompatible with an account of the forgetting effects based on output interference. The output interference account would predict poorer performance on the RP− items whenever there was a robust advantage for the RP+ items, because output interference is caused by prior recall of stronger items, but this was not seen for faces in Experiment 2. In fact, collapsed across cue type, the amount of inhibition (the difference between recall for RP− and U items) was unrelated to recall performance for RP+ items in both experiments ($r = 0.10$ for Experiment 1, and $r = .10$ for Experiment 2). This is not due to restricted range associated with ceiling effects in the RP+ scores, since no correlation emerges between RP+ and inhibition even when those close to the ceiling are excluded. For Experiment 1, when those recalling seven or more RP+ items are excluded, the correlation between RP+ and inhibition for the remaining 52 participants is $r = 0.06$. For Experiment 2, the same exclusion criterion leads to a correlation of $r = -.07$ across 23 participants. In addition, comparison across the conditions involving face cues (either singly or in combination with the category cues) reveals a pattern that is incompatible with the output interference explanation. In those conditions, the presence of a face specifies the output order, and so output interference is matched. However, whilst neither of the single face-cued conditions shows a retrieval-induced forgetting effect, the joint cue condition in Experiment 2 does.

An additional concern with the studies, which complicate interpretation of the retrieval cues is that the faces themselves were presented at the same time as the categories and exemplars. Thus, as well as forming associations between the face and exemplars, participants will also have been forming associations between the faces and the category cues. Thus, it might be argued that the face cues were not truly independent of the category cues.

In addition, it might be argued that since all stimuli were originally encoded with a face, it is possible that all items are now similar. That is, since faces are confusable, all items are now associated with the over-arching category of being associated with a face. Thus, at practice, as well as retrieving a subset of the practiced category (e.g., SPORT), participants are also retrieving a sub-set of items that were associated with a face. This would therefore predict that all items, including baseline items, might be subject to retrieval inhibition, and thus the baseline against which inhibition is being measured may be inappropriate.

Both these potential criticisms stem from the use of faces. Whilst it is not easy to see how these criticisms can fully explain the patterns observed in the previous experiments, it seemed sensible to conduct an experiment that responds to these two points. In the experiment that follows, the independent cues were unrelated words, which were independent of the category cues both semantically, and episodically. The only link between the independent cues and the exemplars was an episodic association created immediately prior to the retrieval-induced forgetting experiment.

Experiment 3

Experiment 3 was therefore designed to meet the concerns outlined above. Instead of using faces as independent cues, words with no association with any items were used as paired associates of each category exemplar. Thus, there was no uncertainty as to which episodic cue was paired with which target word. To remove the direct episodic association between the episodic cues and the category cues, participants first learned the episodic cue—exemplar association, before going on to learn the category—exemplar association, i.e., an A–B, C–B design was employed. Finally, at test, we ensured that we tested recall for all RP− items prior to recall of U items, which in turn preceded recall of RP+ items to control for output interference.

If the previous cue-dependent effects were due to any of these factors then cue-independent forgetting would be predicted in this final study. On the other hand, should cue-dependent effects be observed once again, it would add weight to cue-specific interpretation of those data.

Participants

A total of 72 participants were tested in this study. All were either undergraduate or postgraduate students recruited via posters distributed around the University of Plymouth. None of them had participated in the previous studies. Participants were randomly assigned to one of two experimental conditions. No other biographical data were recorded.

Procedure

This experiment consisted of four phases. Initially there was a learning phase, in which participants learned the associations between the exemplars and the to-become episodic cues. Participants saw 48 word pairs (e.g., *zinc—apple*), in random order, at a rate of 4 s per pair, with a 1 s inter-stimulus interval, presented on screen in 44 point Times Roman font, using a PowerPoint slide show. They were told to remember the words, such that when they next see the cue word they would be able to recall the target word associated with it. Following presentation of the entire list participants were given a cued-recall test for all items, in a different random order. The cue words appeared on the screen as before, and participants responded verbally. Participants then repeated this procedure a second time, using a different random order for presentation and test, but with a faster rate of presentation (2 seconds per pair). All other aspects of presentation were the same. Performance on the list was scored after each trial. If participants successfully recalled 60% of items the learning phase terminated. If they did not, they repeated the learning—test cycle again, using a 2 s presentation rate. If they did not achieve 60% on the third test they were dropped from the remainder of the study and replaced. On average, participants recalled 81.2% of the items by the end of their final study list.

The second phase of the study was the standard learning phase of the retrieval-induced forgetting paradigm, used previously. Category cue—exemplar pairs were presented on screen at a rate of 3 s per pair, with a 1 s inter-stimulus interval. Six categories were used, with eight exemplars per category. These were presented in random order, and participants were told to try to remember the word pairs. They were told that the target words would be the same as they had seen previously, but they were not explicitly told to associate responses in this stage with the previous one. Following presentation of the entire list, participants engaged in retrieval practice for half the items from 2/3 the categories. There was one deviation from the procedure used in the category-practice conditions of the prior experiments. Pilot work revealed that standard category—letter stem cues at test produced ceiling effects in all conditions. Consequently we used category plus fragment cues as explicit memory cues at final test and practice (e.g., *FRUIT*—_ _ *p* _ _).

There then followed a 7 min filled delay prior to a final recall test for all items. Half the participants received category cues, whilst the other half received the episodic cues that had originally been paired with the target words. Participants were randomly assigned to these two conditions. The two groups did not differ in their degree of learning of the original episodic—target word pairs, $t(70) = 1.45$ with those having being tested by category cues having originally recalled 83.3% of the items, and those tested by episodic cues having recalled 79.1%. As described above, we tested using fragments plus categories, so that we could fix output order in that condition. This enabled us to test recall for RP− items first, followed by U items then RP+ items. The use of fragments for the episodic cues was not necessary, since each episodic cue was associated with a single target word. The same test order was maintained for both lists.

Results and discussion

The data were scored and analyzed as in the previous two experiments. A 3 (item status: RP+, RP, and U) × 2 (condition: category cue, unrelated cue) ANOVA revealed main effects of item status, $F(2, 140) = 40.5$, $MSE = 116.9$, $p < .001$, and condition, $F(1, 70) = 52.4$, $MSE = 868.0$ $p < .001$, and a significant interaction between the two, $F(2, 140) = 37.7$, $MSE = 109.1$, $p < .001$. As in the previous experiments, the follow up analyses focused on the retrieval-induced forgetting effect in each condition. Following retrieval practice, participants were cued with either category cues (i.e., standard retrieval-induced forgetting paradigm) or with episodic cues. For the category cues, there was a reliable retrieval-induced forgetting effect, $t(35) = 3.33$, as well as a reliable retrieval practice effect, $t(35) = 7.35$. However, for the episodic cues, there was no retrieval-induced forgetting effect, or practice effect, $t < 1$ in both cases. Table 4 gives the mean performance for RP+, RP−, and U items in each condition.

Quite clearly, performance was superior with the episodic cues, for all kinds of items. In itself, this is not meaningful, since the absolute level of performance from the episodic cues is a function of the degree of initial learning. Presumably, if participants had received less experience with the initial episodic cues, or if we had increased the delay between that study phase and the category—exemplar study phase, they would have shown worse performance with those cues. However, since our aim was to use the episodic associates to cue memories that would be hard to remember, our method favored a high degree of initial learning. Nonetheless, it remains a possibility that the high level of performance is masking an inhibitory effect on the RP− items. To explore this issue, we took the same strategy as we had used previously, and sought to match performance on the U items by excluding the poorest performing participants in the category cued condition, and the best performing participants, based on their recall of U items. The results are shown in the bottom section of Table 4. Here both groups recall the same level of U items, $t(33) < 1$. However, whilst the forgetting effect is still seen in the category cued condition, $t(16) = 3.08$, it remains absent in the episodically cued condition, $t(17) = 1.03$. Moreover, for this sample matched on recall of U items, there is no difference in recall of RP+ items, $t(33) = 1.00$, but the category cued sub-sample recall significantly fewer RP− items, $t(35) = 2.13$. Thus, as in the previous two experiments, the retrieval-induced forgetting effect in this study is limited to the cues that were used to retrieve the competitor item, and does not appear to be a baseline artifact.

General discussion

Our aim was to test the current inhibitory account of retrieval-induced forgetting (Anderson, 2003; Anderson & Spellman, 1995; Levy & Anderson, 2002) using a new independent probe technique. Previous studies which have demonstrated forgetting using an independent probe at the final test stage have used pre-existing semantic relations between the target items and the independent probes (e.g., Anderson et al., 2000; Anderson &

Table 4
Mean (and *SD*) number of items recalled for practiced items from practiced categories (RP+), unpracticed items from practiced categories (RP−), and unpracticed items from unpracticed categories (U), dependent upon retrieval cue used, for Experiment 3

Retrieval cues	Item status					
	RP+		RP−		U	
	Mean	*SD*	Mean	*SD*	Mean	*SD*
Whole sample						
Category	12.42	2.15	7.58	3.41	8.86	3.03
Episodic	13.69	2.27	13.63	2.65	13.55	2.64
Matched on baseline (U) items						
Category (23)	13.17	2.41	9.76	3.13	11.58	1.58
Episodic (14)	12.16	2.71	11.89	2.74	11.33	1.94

Notes. Category, category cue used at retrieval. Episodic, episodic cue used at retrieval. Numbers in parentheses represent the number of participants in each condition, matched on performance on U items.

Spellman, 1995). As discussed above, we have some reservations about this approach. Instead we provided participants with two potential cues at encoding. One of these cues was then used to elicit a retrieval competition, but either, or both, of the cues could be used as a final test cue. Thus, our independent probes were episodically related to the unpracticed items, rather than semantically related. However, the logic of our approach was identical to that used by Anderson and colleagues (e.g., Anderson et al., 2000; Anderson & Spellman, 1995). If retrieval practice leads to inhibition of the related item then forgetting should be seen whatever the retrieval cue used at test, whether semantically related, or episodically related.

The results of this alternative methodology appear to be at odds with current theorizing about retrieval-induced forgetting. We obtained the standard retrieval-induced forgetting effect when we used the same retrieval cue practice and test. However, the key test for an inhibitory account of this phenomenon is forgetting with independent cues. In all three studies, we failed to see retrieval-induced forgetting when our independent episodic cues were used. Additionally, across the two studies involving faces, retrieval-induced forgetting was only seen for the joint face plus category cues in the final test when joint cues were used at practice.

Comparison of the relative retrieval practice benefits (improved recall of RP+ items) and costs (worse recall of RP− items) are particularly informative on two counts. First is that the data across the three studies clearly indicate that the lack of a retrieval-induced forgetting effect is not simply due to a lack of a retrieval practice effect. In Experiments 1 and 3, the episodic cues produced neither retrieval-induced forgetting, nor benefits of retrieval practice. However, in Experiment 2, the face cues did reveal a retrieval practice effect, but the retrieval inhibition effect was still absent. At the same time, the joint cues revealed robust retrieval practice effects in Experiments 1 and 2, yet only showed a forgetting effect in Experiment 2. Thus, the pattern of effects on retrieval-induced forgetting cannot be explained away by claiming a lack of a retrieval practice benefit as a causal mechanism.

As well as ruling out simplistic accounts of our lack of forgetting effects in some conditions, these data also say something rather fundamental about the nature of the retrieval tests, and the assumptions made in the current theoretical modeling of these effects. The standard retrieval-induced forgetting paradigm requires retrieval practice for a subset of items. According to Anderson and Spellman's (1995) feature competition account this results in activation of features of the practiced items and inhibition of features of competitor items, compared to baseline. This pattern of activation and inhibition is argued to be detectable with an independent probe. However, in the present work, not only are the retrieval costs not detectable with an independent cue, in some cases the retrieval benefits are not apparent either. In Experiments 1 and 3, repeatedly retrieving an item to a category cue led to much better recall if the same cue was used during the final test (either singly or in combination with the faces), but had no impact on the subsequent retrieval of that item to a face cue alone. This is not compatible with the idea that there is a single representation of an item that is accessed by both cues.

But, if different cues do not access the same representations where does this leave the independent cue technique? Clearly, it strongly suggests that the episodic representations that are activated during memory studies are not context-independent, but rather that they are multifaceted, and different retrieval cues access different aspects of those events. This is not a novel observation. It is essentially a description of Tulving and Thomson's (1973) encoding specificity principle, with two additional assumptions. One is that the retrieval practice phase constitutes an encoding opportunity, which is well established (Landauer & Bjork, 1978). The second assumption is that the result of the second retrieval opportunity is strengthening of associations between the specific retrieval cue and those aspects of the event that were retrieved. As a consequence, the strength of the association between the retrieval cue and items previously associated with that specific cue is reduced, relative to the practiced cue. This creates transfer appropriate forgetting.

This account can readily account for the current set of data. In Experiment 1, the retrieval competition is induced by category cues, and so when faces are used as retrieval cues, either singly or in combination, then retrieval-induced forgetting is absent. Likewise in Experiment 3, the unrelated words are unrelated to the cues used to elicit the competition in the practice phase. In Experiment 2, when both categories and faces are present at the competition then retrieval-induced forgetting is seen when joint cues are used at test, as well as when the category cue is used at test. One difficulty with this account is that the faces that were present with the RP+ items are different from those used as retrieval cues for the RP− and U items. Thus, it could be argued that the particular retrieval competition that was run during retrieval practice was not replicated at final test, and so retrieval-induced forgetting effects should be absent. However, it is also reasonable to argue that at an abstract level there is considerably more overlap between the practice and test phases in Experiment 2 than in Experiments 1 and 3. That is, in Experiment 2 retrieval from a category cue occurs in the presence of an unfamiliar face for both practice and test phases. To the extent that faces share some abstract level of representation, the practice and final test stages were similar.

How could this view account for the cue-independent effects reported previously using semantically related independent cues? Notwithstanding the reservations already expressed, it might be argued, in line with the above, that those studies had different practice and final test cues, and so retrieval-induced forgetting should be absent. We would argue instead that the retrieval competition at practice and the final test are in fact quite similar in the previously published studies using the independent probe technique. The semantically related independent cues are by their nature related to both the target, and the category cue. If they were not, they could not be used as retrieval cues. Consider Anderson et al. (2000) independent cue methodology. Participants studied the category RED with the following items: *brick*, *heart*, *fire*, *sunburn*, *tomato*, *apple*, *cherry*, and *radish*. Following retrieval practice, memory was tested either with original cue (RED) or the "independent" cue FOOD. But, in a context where half the red items are foods, and no other foods are present, how different are the cues RED and FOOD in functional terms? One could easily see how participants might closely link the concepts of red and food during the study phase, so that whichever retrieval cue is used, the effects are highly similar. This is quite different from the conditions of the present study in which there is no a priori association between the category cues and the independent episodic cues.

What causes context-specific retrieval-induced forgetting?

Up to this point, we have been deliberately theoretically neutral as to the mechanism by which the context-specific retrieval-induced forgetting effect (transfer appropriate forgetting) occurs. There appear to us to be two possibilities, both of which are context specific effects. The first is a relatively small step from the position advocated by Anderson and Spellman (1995), namely that the losers of a retrieval competition are inhibited as a result. The only point of departure from Anderson and Spellman's (1995) account would be that the inhibitory effect is modulated by context. There are two ways of expressing this view. One is that since a context-specific representation is accessed during retrieval, it is a context specific representation that is inhibited, rather than any general representation of a concept. A slightly different account would be that it is the general concept that is inhibited, but only within a specific retrieval context.

The potential contribution of context to inhibitory phenomena has been powerfully demonstrated recently by Tipper, Grison, and Kessler (2003) in the inhibition of return paradigm. They conducted a series of standard inhibition of return trials, each superimposed on a different contextual background, which consisted of different photographs of faces. In contrast to standard inhibitory effects of a few seconds, or trials, Tipper et al. (2003) found inhibition of return effects lasting up to 13 min, spanning up to 192 trials. Whilst the authors maintain that inhibitory processes are at work in their paradigm, it is clear that they do not regard those processes as existing continuously over time, but rather as being contextually reinstated at test. Since the location to-be-inhibited is involved in numerous trials, the inhibition cannot be permanently active. Instead, it is reinstantiated when the faces present on the prime trial are represented at test. Their argument that, "the existence of this long-term inhibitory effect relies on the retrieval of the specific prior processing episode" (p. 24) would serve equally well as a description of the results of the present work.

However, the starting assumption that context specific representations are being accessed offers the possibility of re-evaluating the evidence concerning the non-inhibitory accounts of retrieval-induced forgetting in the first instance. It may be possible to account for the pattern of findings reported to date using a modified associative model. Whilst it is beyond the scope of the present paper to offer a full account of such a model, we provide a brief sketch of how such a model might account for the present set of data.

Fig. 1 is modeled after the diagrams used in previous papers discussing associative and inhibitory accounts of retrieval-induced forgetting (e.g., Anderson, 2003; Anderson & Spellman, 1995). However, in addition to the hypothetical associative links between the categories and exemplars are some additional links with context. There are two kinds of context represented in the figure. At the specific item-level, there is an individual episodic context associated with each target item, which corresponds to the faces used in the first two experiments, and the unrelated words in the final study. More generally, there is the experimental context itself, which is divided into the different phases of the experiment, the study context, a retrieval practice context and a final test context. Of course, these are embedded within an overall experimental context, which itself will be embedded within a broader context of the participants ongoing activities. The assumption is, all other things being equal, context drifts over time, such that adjacent events share more context than more distant events. (e.g., Burgess & Hitch, 1992). Thus, in the absence of any other information at test, the most likely responses are those made mode recently, namely the practiced items at test.

No doubt there are many other associations learned in these studies, such as inter-item associations, and associations between category cues and episodic cues. However, rather than include all possible links, a simple model is sketched out. In reality it is anticipated that activation would cascade through such a system, and that at all levels of item, category and context, there are distributed representations.

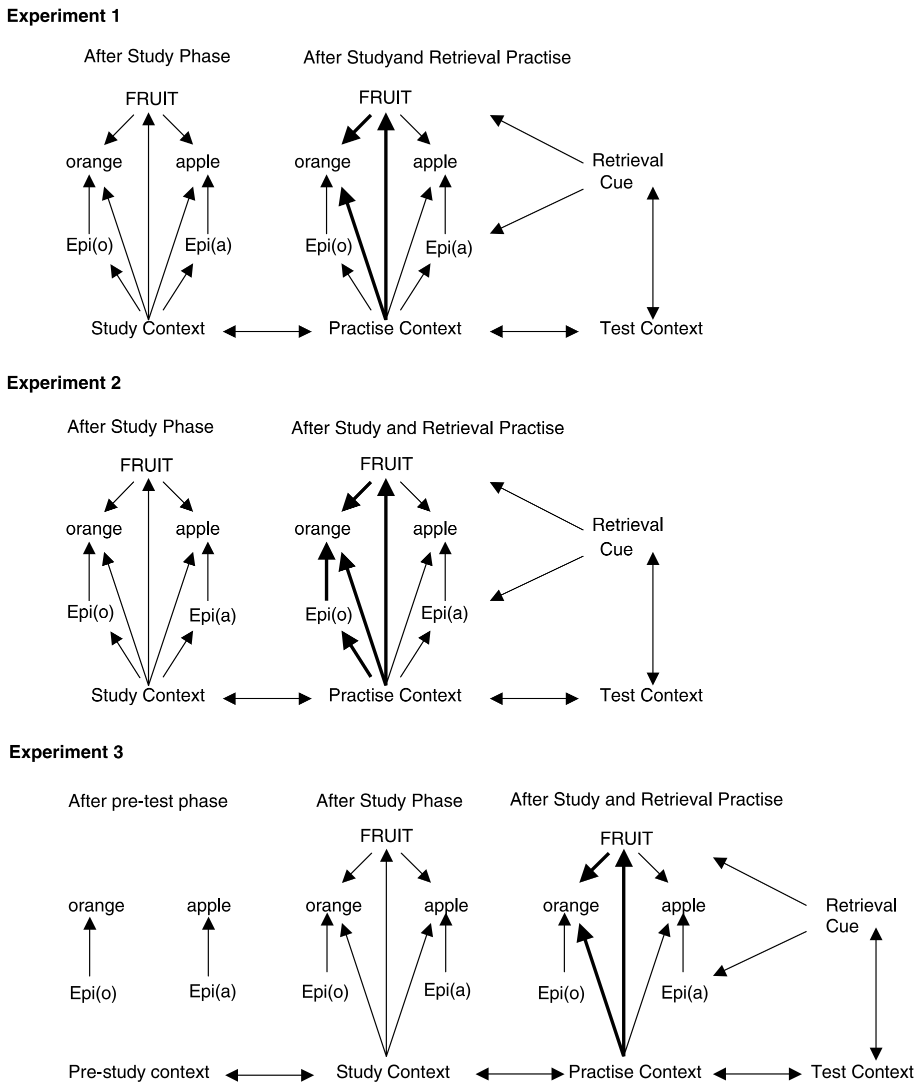

Fig. 1. An associative model of the cue-dependent retrieval-induced forgetting effects seen in the present studies. *Notes:* Epi (o) = Episodic cue associated with *orange*. Epi (a) = Episodic cue associated with *apple*.

After encoding, each of two items (*apple* and *orange*) are equally well learned, and each equally well associated with the category cue (FRUIT), their specific episodic cue (a face) and the study context. This is represented in Fig. 1 by the equivalent strengths of line between all elements. Later, during retrieval practice, the item *orange* is repeatedly retrieved to the cue FRUIT. This has two effects. First the association between FRUIT and *orange* in memory is strengthened (as represented by the stronger line connecting the two in Fig. 1). At the same time, the cue FRUIT now becomes associated with the retrieval practice context, as does the item itself. The model assumes that these effects are additive to the previously learned associations, with nothing happens to the association between fruit and apple, i.e., there is no inhibition in this model.

At test, participants are given one of three retrieval cues. If they are given the category cue FRUIT—a____?,This will lead to the direct activation of the cue FRUIT, which has a strengthened association to the item *orange.* At the same time, the similarity between the test context (FRUIT—*a* _____) and practice context (FRUIT—*or* ____) may also lead to activation spreading from the contextual representation to the item *orange.* Thus, the item orange is likely to receive a retrieval advantage to the cue fruit, and so occlude the retrieval of apple. Now consider what happens when the participant is presented with the episodic cue previously associated with *apple*. This will directly access the prior occurrence of the episodic cue, and so access apple without activating orange. Since the retrieval cue now does not resemble the practice phase, the effects of context are likely to be weaker, and so again, no benefit to orange will accrue. Thus, no retrieval-induced forgetting is predicted. Finally, consider the joint cue. Here there is a direct access to the episodic cue to *apple*, as before, but some contextual similarity to the practice phase. However, this similarity is low since a combined cue of category plus face is rather different from a category cue alone. Thus, little, if any retrieval-induced forgetting is expected.

For Experiment 2, the main difference in the design is that the episodic cue for orange was present at retrieval. Thus, following retrieval practice, the links between context and the face and target are both strengthened, as well as the category-target association. Now, consider the effects of the three kinds of retrieval cue. The category cue, as before, will have a strengthened association with orange, and so will predict forgetting of apple due to occlusion. However, this effect may be moderated by a weakened association between the current test context (a category alone) and the practice context (face plus category). Also as before, the face cue can directly access the item *apple.* This effect may also be moderated by the slight increase in contextual similarity between the test context (a face) and the practice context (context plus face). Finally, for the joint cues, as before the category cue is likely to favor the practiced item (orange) whilst the episodic cue should favor the target item (apple). However, the similarity between test and practice context is now stronger than Experiment 1 (since both involve a category plus a face), and so a stronger contextual effect is predicted. Since the prior context was associated with orange, this may also favor recall or orange. Thus, compared to Experiment 1, there is more bias towards orange in Experiment 2, and so more occlusion of apple. Thus, Experiment 2 shows retrieval-induced forgetting for the joint cues whilst Experiment 1 does not.

The account for Experiment 3 is essentially the same as Experiment 1, except that the subjects associate each target word to an episodic context that is temporally prior to the presentation of the category cues. Thus, at study, and practice, there is no association between those contexts and the episodic cues. Otherwise the situation at the end of the practice phase resembles Experiment 1. With category cues, there is both a direct bias towards *orange* and away from *apple*, and a contextual bias in the same direction. However, with the episodic cue, there is a direct association with *apple*, and no contextual bias at all. Hence no retrieval-induced forgetting is predicted. In fact, one might even also wish to additionally assume that the test context with episodic cues most closely resembles the test phase during the pre-learning phase, and so that too might support access to apple, with no consequent expectation of retrieval-induced forgetting effects.

As we acknowledged at the outset, these models are at best outline sketches. They contain any number of assumptions (such as the degree of contextual overlap, the relative effects of strengthening of associations at each stage, and the decay of those associations). However, in principle, they differ little in level of detail from those used to advance the inhibitory account (e.g., Anderson & Spellman, 1995; Levy & Anderson, 2002). Where they do differ is in the role played by context in both the remembering, and forgetting of items in the retrieval-induced forgetting paradigm.

One interesting issue that arises from this model is the question of the role of retrieval in producing the retrieval-induced forgetting effect. It has been repeatedly shown that recall is a necessary prerequisite for the forgetting effect (Anderson et al., 1994, 2000). At face value, this might suggest that the associative model might predict forgetting even if practice is recognition based, since recognition would strengthen one competitor to the exclusion of another. However, the effect could work an entirely different way: recall itself could constitute a contextual cue. That is, at the final recall test, participants are contextually cued to reinstate their last attempt at recall from the same cue. Thus, recall may be necessary for the effect, but not because it produces competition that results in inhibition at practice, but because it offers a powerful contextual cue at practice that is reinstated at the test phase.

Conclusions

Three studies have used a novel independent probe technique to demonstrate that retrieval-induced forgetting can be cue-dependent. This has led us to the notion of transfer appropriate forgetting. In this final section, we would like to reconsider a favorite example that has been used when discussing the need for inhibition in memory retrieval. This is the question of remembering where one parked one's car in the morning (see Anderson, 2001; Anderson & Neely, 1996; Anderson et al., 2000; Bjork & Bjork, 1996; Levy & Anderson, 2002; Macrae & MacLeod, 1999). Trying to recall today's parking spot requires that we do not recall yesterday's parking spot, a task that appears particularly troublesome to some memory researchers. The implication from previous studies of retrieval-induced forgetting is that the consequence of such retrieval competition is that we inhibit our memory for where we parked yesterday. But does such an idea capture the flexibility of mental life? What if yesterday I had to drop the kids off at school, and so was late. As a result I couldn't find a parking place in my regular parking lot, and so parked a long way away from work, forcing me to be even later. However, today I arrived early and parked closer to work in the usual place. All things being equal it may be the case that trying to recall where my car is today makes me less likely to remember yesterday's parking place. That is, if given the simple retrieval cue "parking place today" I may show later forgetting to the cue "parking place yesterday." But what if I recall the school trip, or the fact that I was late to work prior to this final test? Given such a retrieval cue, it seems entirely likely that I will remember where I parked on the day I was late. Given the ubiquitous role of context and retrieval cues in remembering, it seems entirely plausible that there is a similar role for context in (retrieval induced) forgetting.

Appendix. Verbal materials used in the experiments

Experiments 1 and 2

Category cues	Exemplars
FOOD	Cheese, Fish, Pasta, and Beefburger
PET	Hamster, Budgie, Gerbil, and Rabbit
SPORT	Rugby, Tennis, Swimming, and Hockey
DRINK	Tequila, Vodka, Wine, and Lager
HOBBY	Cooking, Reading, Painting, and Drawing
COUNTRY	Spain, Italy, France, and Greece

Experiment 3

Category cues	Exemplars	Unrelated cues
FRUIT	Apple	Zinc
	Orange	Nylon
	Banana	Son
	Lemon	Belt
	Grape	Kidney
	Strawberry	Oil
	Tangerine	Rope
	Mango	Sword
TOOLS	Hammer	Dress
	Pliers	Spider
	Wrench	Cabbage
	Drill	Train
	Ruler	Glove
	Axe	Chair
	Level	Bone
	Bolts	Winter
ANIMALS	Fox	Tulip
	Zebra	Tree
	Lion	Pencil
	Elephant	Ocean
	Cat	Desk
	Tiger	Hand
	Pig	Glass
	Deer	Phone
PROFESSION	Doctor	Sky
	Lawyer	Butter

Appendix (*continued*)

Category cues	Exemplars	Unrelated cues
	Accountant	Ship
	Engineer	Brush
	Carpenter	Piano
	Farmer	Carpet
	Nurse	Shed
	Banker	Wine
LITERATURE	Magazine	Canoe
	Textbook	Bridge
	Newspaper	Year
	Dictionary	Snow
	Comic	Bed
	Journal	Coffee
	Article	Rose
	Letter	Curtain
SPORT	Tennis	Milk
	Fencing	Guitar
	Hockey	Closet
	Cricket	Emerald
	Golf	Tin
	Wrestling	Foot
	Volleyball	Ring
	Rugby	Car

References

Anderson, M. C. (2001). Active forgetting: Evidence for functional inhibition as a source of memory failure. *Journal of Aggression, Maltreatment and Trauma, 4*, 183–210.

Anderson, M. C. (2003). Rethinking interference theory: Executive control and the mechanisms of forgetting. *Journal of Memory and Language, 49*, 415–445.

Anderson, M. C., & Bell, T. (2001). Forgetting our facts: The role of inhibitory processes in the loss of propositional knowledge. *Journal of Experimental Psychology: General, 130*, 544–570.

Anderson, M. C., & Bjork, R. A. (1994). Mechanisms of inhibition in long-term memory: A new taxonomy. In D. Dagenbach & T. H. Carr (Eds.), *Inhibitory processes in attention, memory and language* (pp. 265–325). San Diego: Academic Press.

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., 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., Green, C., & McCulloch, K. C. (2000). Similarity and inhibition in long-term memory: Evidence for a two-factor theory. *Journal of Experimental Psychology: Learning, Memory and Cognition, 26*, 1141–1159.

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., & Neely, J. H. (1996). Interference and inhibiton in memory retrieval. In E. L. Bjork & R. A. Bjork (Eds.), *Memory. Handbook of perception and cogniton* (2nd ed., pp. 237–313). San Diego, CA: Academic Press.

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

Bjork, E. L., & Bjork, R. A. (1996). Continuing influences of to-be-forgotten information. *Consciousness and Cognition, 5*, 176–196.

Burgess, N., & Hitch, G. J. (1992). Toward a network model of the articulatory loop. *Journal of Memory and Language, 31*, 429–460.

Ciranni, M. A., & Shimamura, A. P. (1999). Retrieval induced forgetting in episodic memory. *Journal of Experimental Psychology: Learning, Memory and Cognition, 25*, 1403–1414.

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. Gruneberg, P. E. Morris, & R. N. Sykes (Eds.), *Practical aspects of memory* (pp. 625–632). New York: Academic Press.

Levy, B. J., & Anderson, M. C. (2002). Inhibitory processes and the control of memory retrieval. *Trends in Cognitive Sciences, 6*, 299–305.

MacLeod, M. (2001). Retrieval induced forgetting in eyewitness memory: Gorgetting as a consequence of remembering. *Applied Cognitive Psychology, 15*, 1–15.

MacLeod, M. D., & Macae, 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.

McClelland, J. L., Rumelhart, D. E. (1986). *Parallel distributed processing* (Vol. 2). MIT Press, Cambridge, MA.

Melton, A. W., & Irwin, J. M. (1940). The influence of degree of interpolated learning on retroactive inhibition and the overt transfer of specific responses. *American Journal of Psychology, 53*, 173–203.

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

Smith, R. E., & Hunt, R. R. (2000). The influence of distinctive processing on retrieval-induced forgetting. *Memory & Cognition, 28*, 503–508.

Thomson, D. M., & Tulving, E. (1970). Associative encoding and retrieval: Weak and strong cues. *Journal of Experimental Psychology: General, 86*, 255–262.

Tipper, S. P., Grison, S., & Kessler, K. (2003). Long-term inhibition of the return of attention. *Psychological Science, 14*, 19–25.

Tulving, E., & Thomson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review, 80*, 352–373.

Williams, C. C., & Zacks, R. T. (2001). Is retrieval induced forgetting an inhibitory process? *American Journal of Psychology, 114*, 329–354.