

Retrieval-induced forgetting in implicit memory tests: The role of test awareness

GINO CAMP, DIANE PECHER, and HENK G. SCHMIDT
Erasmus University Rotterdam, Rotterdam, The Netherlands

Retrieval practice with particular items in memory may result in decreased recall of different, semantically related, items. This retrieval-induced forgetting effect has been demonstrated in studies using explicit memory tests. Anderson and Spellman (1995) have attributed retrieval-induced forgetting to inhibitory mechanisms. This hypothesis predicts similar effects in implicit memory tasks. In our first experiment, using Anderson and Spellman's original paradigm, retrieval-induced forgetting was found using an explicit memory test with independent extralist retrieval cues. In our second experiment, using the same materials, retrieval-induced forgetting was also found using an implicit memory test with independent extralist retrieval cues, but only for participants who were aware of the relationship between the study and practice phase on the one hand, and the test phase of the experiment on the other. Thus, test awareness seems to mediate retrieval-induced forgetting in implicit memory tasks.

A number of studies using the retrieval-practice paradigm have shown that inhibition can play a role in retrieval from memory (for a review, see Anderson, 2003). In the retrieval-practice paradigm, participants first study a number of category-exemplar pairs, followed by retrieval practice with half of the studied exemplars from half of the studied categories. Retrieval practice is carried out using a category-plus-word-stem cued recall task. Their memory is then tested for both practiced and unpracticed exemplars. Typically, results show an increase in recall for practiced items (RP+ items) and a decrease in recall for unpracticed items (RP- items) that belong to the same category as the practiced items, when compared with exemplars from categories that received no retrieval practice (NRP items). In this final recall test, the names of the studied categories are used as cues.

Anderson and Spellman (1995) adapted the retrieval-practice paradigm and found retrieval-induced forgetting using independent extralist retrieval cues. These results led Anderson and Spellman to conclude that retrieval-induced forgetting is an inhibitory effect and that the relation between the category and the exemplar or between the practiced and unpracticed items within the category is not inhibited, but rather the unpracticed item itself. They considered cue-independent forgetting to be an empirical criterion for inhibition.

However, if retrieval-induced forgetting is due to item inhibition, it should also be found using different types of tests. In principle, the retrieval-induced forgetting ef-

fect should be found with any type of test assessing the activation of the inhibited item. A number of studies have investigated the validity of this claim using implicit memory tasks rather than the category-cued recall task used traditionally in the test phase of the retrieval-practice paradigm. Unlike what is done in explicit memory tasks such as category-cued recall, no reference is made to the study phase in implicit memory tasks.

Butler, Williams, Zacks, and Maki (2001) used a word-fragment completion task to test implicit memory in the retrieval-practice paradigm. They did not obtain retrieval-induced forgetting. Nor did they obtain retrieval-induced forgetting in several experiments with explicit tests of memory such as word-fragment-cued recall, category-plus-word-fragment-cued recall, and category-plus-stem-cued recall. They found retrieval-induced forgetting only in the standard category-cued recall condition. Therefore, retrieval-induced forgetting seems to depend more on visually presenting a part of the target item than on whether the final memory test is implicit or explicit. In addition, Butler et al. note that most participants detected the connection between the test phase and the study and practice phases. Therefore, the word-fragment completion task could not be regarded as truly implicit.

Perfect, Moulin, Conway, and Perry (2002) used five different tests of implicit memory in the retrieval-practice paradigm: category generation, category verification, perceptual identification, category-cued perceptual identification, and word-stem completion. The retrieval-induced forgetting effect was found only with the category generation task and the category verification task. Perfect and colleagues argue that the reason for this could be that these two tasks are *conceptual* implicit memory tasks and that others not showing the effect are *perceptual* implicit memory tasks. Conceptual memory tasks rely on semantic processing of the items, whereas performance

We thank Manon Augustus and Nelly van der Meer for their assistance in gathering the data and Neil Mulligan for useful comments on assessing and reducing test awareness. Correspondence should be addressed to G. Camp, Erasmus University Rotterdam, Department of Psychology, P. O. Box 1738, 3000 DR, Rotterdam, The Netherlands (e-mail: camp@fsw.eur.nl).

in perceptual memory tasks relies mostly on the processing of the physical attributes of the items (e.g., Blaxton, 1989). If retrieval-induced forgetting is restricted to conceptual tests of memory, this could explain why Butler et al. (2001) did not find retrieval-induced forgetting using explicit perceptual memory tests. However, Perfect et al. did not use independent cues in the category generation task and the category verification task. Therefore, as Perfect and colleagues indicate in their conclusion, their data cannot rule out a noninhibitory account of the reported retrieval-induced forgetting effects.

Moreover, Veling and van Knippenberg (2004) did find retrieval-induced forgetting with perceptual memory tests, using independent cues. They used a recognition test and a lexical decision task with the category exemplars from the study phase as cues. Longer response latencies were found for RP- items than for NRP items in both tests, indicating retrieval-induced forgetting. Thus, Veling and van Knippenberg showed that retrieval-induced forgetting can be found with perceptual memory tests using an independent cue. Longer response latencies for the RP- items in the lexical decision task also seem to indicate that retrieval-induced forgetting can be found with an implicit test of memory. However, it is unclear whether this test was truly implicit. It is conceivable that participants noticed the fact that they were making lexical decisions about words that were presented earlier in the experiment. The possibility of explicit contamination is always a concern in research using implicit memory tests (e.g., Mulligan, 2002; Mulligan & Hartman, 1996). The fact that participants in the Butler et al. (2001) study detected the relation between the test phase and the study and practice phases indicates that this is indeed possible. Noticing this connection could have triggered explicit retrieval of the category, RP+ items, and NRP items, which could have influenced the response latencies for the RP- items in Veling and van Knippenberg's experiment. For their lexical decision task to have been truly implicit, it would have been crucial that the participants not notice the connection between the test phase and the previous phases of the experiment. Veling and van Knippenberg do not report data on this subject, leaving open the possibility that the lexical decision task was not truly implicit.

The experiments reported in this article were conducted to determine whether inhibition can be found in an implicit test of memory using independent cues, without participants being aware of the connection between the previous phases of the experiment and the test phase.

EXPERIMENT 1

The aim of our first experiment was to replicate the retrieval-induced forgetting effect found by Anderson and Spellman (1995) using new materials. This experiment served as a baseline for Experiment 2, in which we used an implicit memory test for the same materials. Participants studied category-exemplar pairs much like those used by Anderson and Spellman in their Experi-

ments 2-4. Half of the exemplars from each category within these pairs belonged to a hidden category not shown to the participants in the study phase or the retrieval-practice phase. We used the retrieval-practice paradigm to test memory for exemplars from these hidden categories with a category-cued recall test.

It should be noted here that Anderson and Spellman (1995) did not measure category-cued recall for the hidden categories independently in their experiments. They first administered a category-cued recall test for the categories used explicitly in the study phase and the retrieval-practice phase, directly *followed* by a category-cued recall test for the hidden categories. Anderson and Spellman then analyzed the data with and without including the additional items recalled in the hidden category test and found retrieval-induced forgetting in both situations. However, it cannot be ruled out that administering the category-cued recall test for the studied categories influenced the recall on the category-cued recall test for the hidden categories administered subsequently. All items that could be generated in the hidden category-cued recall test were also targets in the studied category-cued recall test administered previously. Therefore, performance on these two tests cannot be considered to be independent. It is unclear whether retrieval-induced forgetting would have been found if the category-cued recall test for the hidden categories had been administered independently with the names of the hidden categories as cues. A number of studies have used independent cues (e.g., Anderson & Bell, 2001; Anderson & Green, 2001), but always with a different paradigm and/or different types of materials. Anderson, Green, and McCulloch (2000) did administer an independent category-cued recall test with the names of the hidden categories as cues and found retrieval-induced forgetting. However, their study phase was quite different from Anderson and Spellman's original.

In the present experiment, the study phase was identical to that of Anderson and Spellman (1995), and only the hidden categories were used as cues in the category-cued recall task. In accordance with the inhibitory account of retrieval-induced forgetting, we hypothesized that the retrieval-induced forgetting effect should also be found for the hidden categories when tested independently.

Method

Participants. The participants were 48 psychology students. Most of them fulfilled a course requirement by participating in the experiment, whereas the others received €10 (about \$11.50 U.S.) for their participation. All participants were proficient speakers of Dutch.

Materials and Design. Twelve categories, consisting of six related pairs, were selected. Each category contained six exemplars. None of the exemplars was a member of any of the other 11 categories. However, within each related pair of categories, three exemplars of each category were also members of a hidden category that was not presented explicitly in the study phase or the retrieval-practice phase of the experiment. Only the six items from the two related categories were exemplars of these hidden categories. See Table 1 for examples. All hidden categories were selected on the basis of category norms (Hudson, 1982). The exemplars of the hid-

Table 1
Example of an Experimental and a Control Category Pair
Used in Experiments 1 and 2

Experimental Category Pair		Control Category Pair	
CIRCUS	RIVER	LOUD	SHARP
RP+ Items			
clown	bank	siren	needle
audience	pebbles	alarm clock	toothpick
trapeze	waterfall	scream	dart*
RP- Items		Similar Items	
<i>panther</i>	<i>crocodile</i>	<i>revolver</i>	<i>dagger</i>
<i>elephant</i>	<i>hippo</i>	<i>grenade</i>	<i>spear</i>
<i>bear</i>	<i>rat</i>	<i>bomb</i>	<i>lance</i>

Note—RP+ items are items from practiced categories that received retrieval practice; RP- items are items from practiced categories that did not receive retrieval practice; similar items are items from unpracticed categories that are members of the same hidden category as the RP- items. No items in the control category pair received retrieval practice. Therefore, certain items in the control category pairs served as controls for the corresponding items in the experimental category pairs. Items in italics were also a member of a hidden category. For items in the experimental category pair, the hidden category was ANIMAL. For items in the control category pair, the hidden category was WEAPON. *In Dutch, this word is only used for an object in games, not for a weapon.

den categories were not among the five most typical category exemplars, in order to prevent a ceiling effect in the implicit memory task used in Experiment 2. Their mean position on a frequency-sorted list for each hidden category was 20.5 ($SD = 15.5$). In selecting the exemplars, we avoided items that had strong associations with other items to prevent the use of retrieval strategies based on this association in the test phase. No exemplars shared the first two letters with another exemplar in the same category or with highly typical nonselected exemplars, to ensure that the cue in the retrieval-practice phase matched only one exemplar. Categories were counterbalanced across conditions.

The design used in this experiment was identical to that used in Anderson and Spellman's (1995) Experiment 2. There were three experimental category pairs and three control category pairs. Table 1 shows one experimental and one control category pair. Experimental category pairs consisted of one category that received retrieval practice (CIRCUS) and one category that did not (RIVER). Of the practiced category, three exemplars received retrieval practice (RP+ items) and three did not (RP- items). The RP- items were always the three exemplars from the hidden category. For example, some participants practiced the items CIRCUS-clown, CIRCUS-audience, and CIRCUS-trapeze from the practiced experimental category CIRCUS (these were the RP+ items) and did not practice the items CIRCUS-panther, CIRCUS-elephant, and CIRCUS-bear (these were the RP- items and members of the hidden category ANIMAL). In the unpracticed experimental category (RIVER), none of the exemplars received retrieval practice, but three of the six items also belonged to the hidden category: RIVER-crocodile, RIVER-hippo, and RIVER-rat all belong to the hidden category ANIMAL. These items are called similar items, because they belong to the same hidden category as the RP- items.

Neither of the categories in the control category pairs received retrieval practice. These categories served as controls for the critical items in the experimental categories. To determine whether the RP- items were truly inhibited, the percentage of correctly recalled RP- items on the hidden category cued recall test was compared with that of correctly recalled items in the control category pairs that corresponded with the RP- items.

Procedure. The participants were tested individually or in groups of up to 5 and they were told that they were participating in an experiment on memory and reasoning that consisted of performing several tasks on the computer. In accordance with the retrieval-

practice paradigm (Anderson & Spellman, 1995), the experiment consisted of four phases: a study phase, a retrieval-practice phase, a distractor phase, and a hidden-category cued recall test.

In the study phase, the participants were told that they would be presented with category-exemplar pairs on the computer screen and that they would see multiple examples from multiple categories. Each pair was shown for 5 sec, and the participants were instructed to remember as many pairs as they could and to use the allotted 5 sec to relate the exemplar to its category. Category-exemplar pairs were presented in random order with the restriction that two items from the same or from related categories were not shown successively. In addition to the 12 experimental categories, 6 filler categories with six exemplars each were also presented in the study phase. These filler categories were selected on the basis of category norms (Hudson, 1982), and exemplars were not among the five most typical examples of their category. The filler categories were unrelated to the experimental categories, and exemplars from the filler categories satisfied the same constraints as did the experimental exemplars. The first and the last three items on the study list were filler category exemplars, to control for recency and primacy effects.

In the retrieval-practice phase, the participants performed retrieval practice with three items of three experimental categories and all six filler categories. The name of the category was presented on the screen, followed by the first two letters of an exemplar and a blank line (e.g., CIRCUS-cl_____). The length of the blank line was held constant to avoid giving cues for word length. The participants were told that they would be shown the names of categories from the previous task and the first two letters of a studied category exemplar. They were given 10 sec to complete the word. Items were presented to the participants in random order. This procedure was repeated twice, so that all items received retrieval practice three times. The first and last three pairs were always exemplars from the filler categories. After the retrieval-practice phase, the participants were given a distractor task in which they were asked to solve a number of puzzles. The content of the puzzles was in no way related to any of the categories or exemplars in the experiment. The distractor task took 20 min.

Finally, in the test phase, the participants were given a category-cued recall test for the hidden categories. They were shown the names of all hidden categories on the screen consecutively and in random order. For each category, they were asked to type six exemplars within 30 sec, after which the next category name appeared. It was

made explicit that the participants had not seen these categories before in the experiment, but they were instructed to use exemplars they had seen previously in the experiment. Thus, the test was an explicit category-cued recall test using independent cues.

Results and Discussion

The retrieval-practice success rate was measured for the experimental categories. Each participant had to retrieve a total of nine exemplars: three exemplars from each of three experimental categories. Each exemplar was retrieved three times, resulting in 27 retrieval-practice trials. The average success rate was 79% ($SD = 11.1$), which is similar to the success rates reported by Anderson and Spellman (1995).

The cued recall data for the final memory test are shown in Table 2. The data from 2 participants were removed because their recall rates were more than 2 standard deviations from the mean. Data from another 2 participants were removed in order to leave the counterbalancing design intact. A paired-samples t test showed that fewer RP- items were recalled in the experimental condition than their corresponding items in the control condition [$t(43) = -2.0, p < .05$], showing retrieval-induced forgetting.¹ These results show that retrieval-induced forgetting can also be found in a category-cued recall test using only the hidden category names as cues.

EXPERIMENT 2

We obtained retrieval-induced forgetting in Experiment 1 using category-cued recall as an explicit memory task. The aim of the second experiment was to determine whether this effect could also be found with an implicit memory task using the same materials and independent cues as in Experiment 1. Test awareness was measured to determine whether the memory task was truly implicit. Inhibitory accounts of retrieval-induced forgetting would predict the same results when an implicit memory task is used.

Method

Participants. The participants were 112 psychology students. Most of them fulfilled a course requirement by participating in the

experiment, whereas the others received €10 (about \$11.50 U.S.) for their participation. All participants were proficient speakers of Dutch. None had participated in Experiment 1.

Materials and Procedure. The materials and procedure were identical to those of Experiment 1, except that an implicit category generation task was used in the test phase. The participants were asked to produce six exemplars of each of the given hidden categories. No reference was made to the study phase, making it an implicit test of memory, again using independent cues. Fourteen filler categories were added to the 6 hidden categories to further obscure the relation between the test phase and the previous phases of the experiment. Thus, the test phase included 20 trials. The participants were first given 4 filler categories and then 1 hidden category, followed by 2 filler categories. This last procedure was repeated until all hidden categories had been presented. Hidden categories and filler categories were randomly selected.

After the category generation task, we assessed whether the test had been truly implicit by administering an awareness questionnaire (e.g., Mulligan, 2002; Mulligan & Hartman, 1996), which assessed whether the participants had been aware of the relation between the two parts of the experiment and whether they had consciously tried to remember the words from the earlier part of the experiment during the test phase.

Results and Discussion

Retrieval-practice success rate for the 27 relevant retrieval-practice trials was measured for the experimental categories. The average success rate was 78% ($SD = 15.4$), which is similar to the success rates in Experiment 1 and those reported by Anderson and Spellman (1995).

Generation percentages are shown in Table 2. Data from 16 participants who indicated that they had consciously tried to remember the words from the earlier part of the experiment during the test phase or who had deliberately not reported items from the earlier part of the experiment during the test phase were removed and replaced, because for these participants it was clear that their awareness of the connection between the two parts of the experiment had affected their responses in the test phase. Removal of participants in such cases is a common approach (see Mulligan, 2002). Our main focus was on the difference between participants who indicated that they had noticed the connection between the two phases of the experiment but indicated that they had not used explicit retrieval strategies (aware participants) and those who had not noticed any connection between the two phases of the experiment at all (unaware participants). Fifty-two participants were aware, and 44 were unaware. To retain complete counterbalancing of retrieval-practice categories, we randomly removed data from 4 unaware participants. This did not influence the reliability of our effects.

A paired-samples t test showed that aware participants generated reliably fewer RP- items than control items [$t(51) = -2.1, p < .05$]. By contrast, there was no difference in generation of RP- items and control items for the unaware participants [$t(39) = .48, n.s.$]. Thus, the results indicate that the retrieval-induced forgetting effect was only found for participants who noticed the connection between the previous phases of the experiment and the test phase, and not for participants who were unaware of this relation.

Table 2

Means and Standard Deviations of the Recall and Generation Percentages of RP- Items and Their Controls in Experiments 1 and 2

Variable	<i>M</i>	<i>SD</i>
Experiment 1		
RP- items	35.4	19.4
Controls for RP- items	41.7	18.9
Difference	-6.30*	
Experiment 2		
RP- items (aware participants)	21.8	14.6
Controls for RP- items (aware participants)	28.6	17.8
Difference	-6.83*	
RP- items (unaware participants)	20.3	12.0
Controls for RP- items (unaware participants)	18.9	12.4
Difference	1.39	

* $p < .05$.

GENERAL DISCUSSION

In Experiment 1, we found retrieval-induced forgetting using an explicit memory test and the names of the hidden categories as independent cues. Thus, retrieval-induced forgetting can be found with Anderson and Spellman's (1995) original paradigm, even when memory for inhibited items is tested independently, using cues that were not presented in the study or retrieval-practice phase of the experiment.

Previous studies have shown retrieval-induced forgetting in implicit memory tasks. Perfect et al. (2002) found retrieval-induced forgetting with two implicit memory tests, but these tests did not employ independent cues. Instead, studied category cues were used. Moreover, Perfect and colleagues did not find retrieval-induced forgetting for a number of different implicit memory tests. Veling and van Knippenberg (2004) found retrieval-induced forgetting using a lexical decision task, but we argue that is unclear whether this test was truly implicit, since it is conceivable that participants noticed the connection between the two phases of their experiment. The present study used independent cues to test memory for the inhibited items and an awareness questionnaire to control for explicit contamination. In Experiment 2, retrieval-induced forgetting was found using an implicit memory test and the same set of independent cues as in Experiment 1, but only for participants who were aware of the connection between the two phases of the experiment. Retrieval-induced forgetting was not found when participants were unaware of this relation.

A possible explanation of these findings is interference. In the retrieval-practice phase, retrieval from memory is guided not only by semantic features, but also by contextual features of the study phase. As a result, these contextual features might become stronger cues for RP+ items. In the test phase, at least for aware participants, these contextual features might again be used as cues. Because these are strong cues for RP+ items, these items have a higher probability of being retrieved and might interfere with the retrieval of items that are similar to RP+ items. This interference results in a lower probability of retrieval of RP- items. For unaware participants, no such interference occurs, since the contextual features from the study phase are not used as cues in the test phase. This could also explain why generation of studied items is lower for unaware participants than for aware participants (see Table 2). Not only might contextual features have caused this interference effect, but aware participants might also have retrieved the categories from the study and retrieval-practice phase when they noticed the connection between the two parts of the experiment. Because of the strong link between RP+ items and the category due to the retrieval-practice phase, it is possible that RP+ items were activated in the test phase, which might have caused interference with the activation of RP- items. These explanations would argue against an inhibitory account of the results in both Experiments 1 and 2.

Alternatively, Anderson (2003) argues that the lack of inhibition in implicit tests may be explained by the assumption that only episodic representations are inhibited. Implicit tests might not tap the activation of episodic representations, thereby masking inhibitory effects. However, Anderson also indicates that it is impossible to distinguish between semantic and episodic representations.

We believe that these experiments help to define the scope of inhibitory effects in memory retrieval. Awareness may be a crucial factor in finding retrieval-induced forgetting. Therefore, experiments on retrieval-induced forgetting should include some measure of participant awareness.

REFERENCES

- ANDERSON, M. C. (2003). Rethinking interference theory: Executive control and the mechanisms of forgetting. *Journal of Memory & 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., & GREEN, C. (2001). Suppressing unwanted memories by executive control. *Nature*, **410**, 131-134.
- 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, & Cognition*, **26**, 1141-1159.
- ANDERSON, M. C., & SPELLMAN, B. A. (1995). On the status of inhibitory mechanisms in cognition: Memory retrieval as a model case. *Psychological Review*, **102**, 68-100.
- BLAXTON, T. A. (1989). Investigating dissociations among memory measures: Support for a transfer-appropriate processing framework. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **15**, 657-668.
- BUTLER, K. M., WILLIAMS, C. C., ZACKS, R. T., & MAKI, R. H. (2001). A limit on retrieval-induced forgetting. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **27**, 1314-1319.
- HUDSON, P. T. W. (1982). *Preliminary category norms for verbal items in 51 categories in Dutch* (Internal report). Katholieke Universiteit Nijmegen, Nijmegen: Vakgroep Psychologische Functieleer.
- MULLIGAN, N. W. (2002). The effects of generation on conceptual implicit memory. *Journal of Memory & Language*, **47**, 327-342.
- MULLIGAN, N. W., & HARTMAN, M. (1996). Divided attention and indirect memory tests. *Memory & Cognition*, **24**, 453-465.
- PERFECT, T. J., MOULIN, C. J. A., CONWAY, M. A., & PERRY, E. (2002). Assessing the inhibitory account of retrieval-induced forgetting with implicit-memory tests. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **28**, 1111-1119.
- VELING, H., & VAN KNIPPENBERG, A. (2004). Remembering can cause inhibition: Retrieval-induced inhibition as cue independent process. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **30**, 315-318.

NOTE

1. We also examined the recall percentages of similar items and their controls for evidence of cross-category inhibition. In Experiment 1, we found no reliable difference between the recall percentages of similar items ($M = 38.9\%$) and their controls [$M = 38.6\%$; $t(43) = .08$, n.s.]. However, in Experiment 2, cross-category inhibition was found, but only for the aware participants. The mean difference in generation percentages between similar items ($M = 18.4\%$) and their controls ($M = 23.5\%$) was significant [$t(51) = -1.69$, $p < .05$, one-tailed]. For unaware participants, the difference between similar items ($M = 18.1\%$) and their controls ($M = 20\%$) was not significant [$t(39) = -.52$, n.s.].