Noncompetitive retrieval practice causes retrieval-induced forgetting in cued recall but not in recognition

Tobias Grundgeiger

Abstract Retrieving a subset of learned items can lead to the forgetting of related items. Such retrieval-induced forgetting (RIF) can be explained by the inhibition of irrelevant items in order to overcome retrieval competition when the target item is retrieved. According to the retrieval inhibition account, such retrieval competition is a necessary condition for RIF. However, research has indicated that noncompetitive retrieval practice can also cause RIF by strengthening cue–item associations. According to the strength-dependent competition account, the strengthened items interfere with the retrieval of weaker items, resulting in impaired recall of weaker items in the final memory test. The aim of this study was to replicate RIF caused by noncompetitive retrieval practice and to determine whether this forgetting is also observed in recognition tests. In the context of RIF, it has been assumed that recognition tests circumvent interference and, therefore, should not be sensitive to forgetting due to strength-dependent competition. However, this has not been empirically tested, and it has been suggested that participants may reinstate learned cues as retrieval aids during the final test. In the present experiments, competitive practice or noncompetitive practice was followed by either final cued-recall tests or recognition tests. In cued-recall tests, RIF was observed in both competitive and noncompetitive conditions. However, in recognition tests, RIF was observed only in the competitive condition and was absent in the noncompetitive condition. The result underscores the contribution of strength-dependent competition to RIF. However, recognition tests seem to be a reliable way of distinguishing between RIF due to retrieval inhibition or strength-dependent competition.

Keywords Inhibition · Interference · Memory retrieval

T. Grundgeiger (✉)
Institute for Human-Computer-Media,
Julius-Maximilians-Universität Würzburg, Oswald-Külpe-Weg,
Geb. 82, 97074 Würzburg, Germany
e-mail: tobias.grundgeiger@uni-wuerzburg.de

Practicing a subset of previously learned items can lead to the forgetting of learned, but not practiced, items. Such forgetting has been frequently demonstrated using the retrieval practice paradigm (Anderson, Bjork, & Bjork, 1994). In this paradigm, participants study category–word pairs (e.g., "WOOD–Cabinet," "WOOD–Bench," "GREEN–Lettuce," etc.). Subsequently, in the so-called retrieval practice phase, half of the studied items from half of the studied categories are practiced by providing a category and a specific word stem (e.g., "WOOD–Ca____"). After a distractor, participants are asked to remember all items. The typical finding is that practiced items ("Cabinet") are remembered better than nonpracticed items of a nonpracticed category ("Lettuce"). This effect is called *retrieval facilitation*. More interesting, nonpracticed items of a practiced category ("Bench") are remembered less well than nonpracticed items of a nonpracticed category ("Lettuce"). This effect is called *retrieval-induced forgetting* (RIF).

A prominent theoretical explanation for RIF is retrieval inhibition (Anderson, 2003; Anderson et al., 1994; Storm & Levy, 2012). According to this view, providing a category and a word stem during the retrieval practice triggers a search that activates all items associated to this category. To overcome this retrieval competition and retrieve the correct item, irrelevant items need to be suppressed, and their memory representations are inhibited. Because of this inhibition, these items are more difficult to retrieve in the final memory test and are remembered worst. Consistent with the retrieval inhibition account, RIF has been found, for example, in category-plus-initial-letter cued-recall tests (Anderson, Bjork, & Bjork, 2000), in word recognition (Hicks & Starns, 2004), and with retrieval cues in the final test that were associated to critical items but not explicitly paired with these critical items during the learning phase (Anderson & Spellman, 1995). For a recent progress report on retrieval inhibition and RIF, see Storm and Levy (2012).

The basic RIF effect can also be explained by strength-based models (e.g., Malmberg & Shiffrin, 2005; Mensink &

Raaijmakers, 1988; Raaijmakers & Shiffrin, 1981; Rundus, 1973). Following such a strength-dependent competition account, the association between a category and the practiced item is strengthened during retrieval practice. In the final memory test, items exhibiting a strengthened association with the respective category interfere with the retrieval of relatively weaker, nonpracticed items. The reduced relative strength leads to impaired recall of the nonpracticed items. For recent critical reviews of key assumptions of the retrieval inhibition account and of the potential explanations arising from a strength-dependent competition account, see Verde (2012) and Raaijmakers and Jakab (2013).

One specific assumption of the retrieval inhibition account is *retrieval specificity* (Anderson, 2003). According to this assumption, RIF should be present only if items have to be actively retrieved during retrieval practice. If a category and a word stem of a learned item are presented, all items belonging to this category are triggered, and inhibitory processes are needed to inhibit competing items and assist in retrieving the correct item. If a target item is, for example, presented intact for restudy during the practice phase, there should be no need for such inhibitory processes. The retrieval specificity assumption has been studied using two different procedures. First, in the *restudy* procedure, items are presented intact during the practice phase, and participants are instructed to relearn the items (e.g., Anderson & Bell, 2001; Bäuml & Aslan, 2004; Ciranni & Shimamura, 1999; Staudigl, Hanslmayr, & Bäuml, 2010). As compared with the standard retrieval practice, restudy resulted in similar improved memory performance for practiced items, as compared with control items, but restudy did not result in RIF. Second, in the *noncompetitive retrieval practice* procedure, participants do not practice the items but practice the corresponding categories. They are presented with the intact word and have to retrieve the corresponding category. As compared with the standard retrieval practice (*competitive retrieval practice*), practicing categories (*noncompetitive retrieval practice*) resulted in improved performance for items that had been used as cues during the practice phase, but noncompetitive retrieval practice did not result in RIF (e.g., Anderson et al., 2000; Hanslmayr, Staudigl, Aslan, & Bäuml, 2010). The above findings have been interpreted to be consistent with the retrieval inhibition account (for an alternative interpretation, see Raaijmakers & Jakab, 2013).

In contrast to the above results, recent studies have demonstrated that RIF can occur without retrieval competition during the retrieval practice phase, thus challenging the retrieval specificity assumption of retrieval inhibition (Jonker & MacLeod, 2012; Raaijmakers & Jakab, 2012). Following the procedure of Anderson et al. (2000), Raaijmakers and Jakab (2012) used noncompetitive retrieval practice and showed participants learned items in the retrieval practice phase,

asking them to retrieve the associated category. In a final cued-recall test, reliable RIF was found. There are several differences between previous noncompetitive retrieval practice studies and Raaijmakers and Jakab's study (for a full discussion, see Raaijmakers & Jakab, 2012). First, retrieval practice in previous studies (e.g., Anderson et al., 2000; Hanslmayr et al., 2010) might have been too easy to strengthen cue–item associations (i.e., by using well-known categories and items such as "FRUIT" and "Orange" and by providing word stems such as "FR___" for retrieval practice). Second, in previous studies, no feedback was given after each trial. Therefore, some category–word associations are strengthened and others might not be strengthened at all in the event that the categories (or words, in the standard retrieval practice procedure) cannot be retrieved. To maximize the strengthening of cue–item associations, Raaijmakers and Jakab (2012) used categories that group items on the basis of properties (e.g., "ROUND–Ball") instead of semantic categories, provided only one first letter or no letter of the asked category, and provided feedback on the correct answer after each trial in the retrieval practice phase.

Raaijmakers and Jakab's (2012) finding is important because it demonstrates that, in the standard retrieval practice paradigm, RIF can be the result of strength-dependent competition. In the absence of retrieval competition, the retrieval inhibition account would not predict RIF in a final memory test. By contrast, a strength-dependent competition account of RIF would predict forgetting, since the association between the category and the item is strengthened during retrieval practice. Being able to separate inhibitory and noninhibitory causes of RIF is, however, important to the advancement of theoretical accounts (Storm & Levy, 2012).

One way of distinguishing between forgetting due to retrieval inhibition and strength-dependent competition might be the use of recognition tests. Indeed, in the context of the retrieval‚practice paradigm, it has been argued that providing only items during the final memory test should render the associative strength between category and item irrelevant (e.g., Aslan & Bäuml, 2010; Gómez-Ariza, Lechuga, Pelegrina, & Bajo, 2005; Hicks & Starns, 2004; Spitzer & Bäuml, 2007; Veling & van Knippenberg, 2004). In other paradigms, in which forgetting is primarily explained by strength-dependent competition, the results are mixed. For the list strength effect, for example, some studies show a list strength effect in recognition (e.g., Buratto & Lamberts, 2008), but others do not (e.g., Kinnell & Dennis, 2011; Malmberg, 2008). However, findings from studies on output interference (e.g., Criss, Malmber, & Shiffrin, 2011) indicate that, in principle, recognition tests are sensitive to strength-dependent competition.

Despite this disagreement, which I will return to in the Discussion section, the contrast between cued-recall tests

and recognition tests has been used to investigate children's inhibitory capabilities in RIF. Aslan and Bäuml (2010) have argued that RIF in cued-recall tests may be caused by inhibition during retrieval practice but also may be due to interference of stronger (i.e., practiced) items during the final memory test. Using both category-cued-recall tests and recognition tests, Aslan and Bäuml (2010) have observed that adults and second-graders showed RIF in both cases. However, kindergarteners showed RIF in the cued-recall test but not in the recognition test. In line with the according literature, this finding suggests that kindergarteners have inefficient inhibitory processes. Therefore, kindergarteners do not show RIF due to retrieval inhibition.

Experiment 1 follows Aslan and Bäuml's (2010) logic by comparing RIF in cued-recall tests and recognition tests, with the aim of investigating the contribution of strength-dependent competition to RIF in younger adults. To this end, competitive and noncompetitive retrieval practice was followed by either category-plus-initial-letter cued-recall tests or old/new recognition tests.

Experiments 2a and 2b aimed to replicate key findings of Experiment 1 and test the potential contribution of reinstating learned cues during the test phase. The contribution of strength-dependent forgetting in recognition tests may be excluded only if participants do not try to reinstate learned cues during the test phase (so called *covert cuing*). Participants, for example, might think back to the learning phase and the learned categories in order to determine whether an item has been studied or not. In line with this idea, researchers have recently made the argument that participants might reinstate these learned cues during the final memory test (e.g., Camp, Pecher, & Schmidt, 2005; Camp, Pecher, Schmidt, & Zeelenberg, 2009; Verde & Perfect, 2011). If this is the case, the influence of strength-dependent competition in recognition tests cannot be excluded. Experiment 2a specifically addresses this argument by presenting the learned category names and encouraging participants to use these learned cues in a recognition test.

Experiment 1

The aim of Experiment 1 was to test the contribution of strength-dependent competition to RIF. To this end, I used a noncompetitive condition (retrieval practice of category names) and a competitive condition (retrieval practice of items). In both conditions, retrieval practice was followed by either a cued-recall test or a recognition test.

For competitive practice and cued recall, both accounts would predict RIF. For the noncompetitive practice and cued recall, only the strength-dependent competition account would predict RIF. For the competitive practice and recognition, the retrieval inhibition account would predict RIF. The

strength-dependent competition account would predict RIF only if participants tried to reinstate learned cues.

For noncompetitive practice and recognition, the retrieval inhibition account would not predict RIF. Again, the strength-dependent competition account would predict RIF only if participants tried to reinstate learned cues. Therefore, any forgetting observed in this condition could be attributed exclusively to strength-dependent competition, which would have become effective due to attempts to reinstate the original cues during the final recognition test.

Method

Participants

A total of 224 university students (175 female) with a mean age of 22 years (range: 18–43 years) participated in exchange for course credit or payment. All participants were native German speakers. Participants were assigned to one of four groups: competitive practice and cued recall ($n = 40$), noncompetitive practice and cued recall ($n = 40$), competitive practice and recognition ($n = 72$), or noncompetitive practice and recognition ($n = 72$). Participants were tested in groups of 1–4.

Design

The factors *retrieval practice situation* (competitive, noncompetitive) and *final test* (cued recall, recognition) were manipulated between participants. The factor *retrieval practice status* was manipulated within participants. For competitive conditions, half of the studied target items from half of the categories were practiced by retrieving the target item itself, given a certain category (Rp+). For noncompetitive conditions, half of the studied target items from half of the categories were practiced by retrieving the category for the given target item (Rp+). In both conditions, the respective other halves of target items were not practiced but belonged to a practiced category (Rp−). The items of the second, nonpracticed half of the categories were divided into items (Nrp+) serving as controls for Rp+ items and items (Nrp−) serving as controls for Rp− items. Finally, the retrieval practice created two types of lure items that appeared as new items only in the recognition test: items from practiced categories (Rp lures) and items from nonpracticed categories (Nrp lures).

Materials

Following Raaijmakers and Jakab (2012), this study attempted to make the category practice challenging by using rather unusual categories and target items with a relatively low taxonomic frequency. Eight categories from existing norms were selected (Mannhaupt, 1983; Van Overschelde, Rawson, & Dunlosky, 2004): "place to live," "means of transportation,"

"a liquid," "building material," "a type of reading material," "green," "wood," and "fly." The German translations of the category names consist of a single word. For each category, six target items with a low-to-medium taxonomic frequency were selected (*Med* = 20). As in previous studies using recognition tests to access RIF (Aslan & Bäuml, 2010, 2011), three frequent exemplars were selected as frequent lures (*Med* = 8) for each category and three exemplars as infrequent lures (*Med* = 29). The six target items in each category were selected to belong to only one category. Within a single category, no two items began with the same initial letter. The target items were between 4 and 12 letters long (*Med* = 6), and lure items were between 3 and 11 letters long (*Med* = 6).

Two category sets consisting of four categories each were created. For each category set, the six items of each category were further divided into two word sets of 4×3 words. Within each of the four experimental groups, the four category–word sets were practiced equally often during the retrieval practice phase. Therefore, all items occurred equally often in each condition (Rp+, Rp−, Nrp+, and Nrp−) across participants, and lures served equally often as Nrp lures and Rp lures.

Procedure

The experimental procedure consisted of three phases: the study phase, the retrieval practice phase, and the final memory test. All instructions were given on the computer screen.

In the study phase, participants were instructed to learn the presented category–word pairs ("WOOD–Cabinet") for an unspecified memory test. A fixation cross was presented for 250 ms, followed by a category–word pair for 3 s. The study list consisted of six blocks. Each block consisted of one randomly selected category–word-exemplar from each of the eight categories, with the restriction that no two items of the same category appeared consecutively at a block border. The study list was presented twice to each participant. Three filler pairs from unrelated categories were presented at the beginning and the end of the study phase. After the study phase, participants solved simple math problems for 1 min as a distractor task.

In the retrieval practice phase, participants in the competitive conditions were presented with a learned category and the word stem of an item ("WOOD–Ca____") and were instructed to retrieve the learned word. Participants in the noncompetitive conditions were presented with a learned item ("____–Cabinet") and were instructed to retrieve the learned category. In both conditions, each trial started with a fixation cross lasting 300, 450, or 600 ms. This duration was determined randomly. Each item was presented for 10 s, and participants were required to type their answer in a box below the item and press the enter key. Each item was followed by a blank screen for 250 ms. The practice list consisted of three blocks. Each block consisted of one randomly selected

exemplar from each of the four practiced categories, with the restriction that no two items of the same category appeared consecutively at a block border. The practice list was presented three times to each participant. The retrieval practice phase was followed by further math problems for 6 min.

In the final memory test, for the cued-recall test, participants were presented with a learned category and the initial letter of an item ("WOOD–C____") and were instructed to retrieve the learned word. As for the retrieval practice, each trial started with a fixation cross (300, 450, or 600 ms), and participants were provided with 10 s to type their answer, followed by a blank screen for 250 ms. Half of the participants within each group started with a practiced category, and the other half with a nonpracticed category. Practiced and nonpracticed categories were tested in alternating order, and the actual categories were selected randomly. Within a category, Rp− or Nrp− items were always tested before Rp+ or Nrp+ items to avoid output interference for Rp− or Nrp−. The order of items within each word type was randomly assigned.

For the recognition test, participants were randomly presented with the 48 learned items and the 48 lure items. Again, each trial started with a fixation cross (300, 450, or 600 ms) before a single item was presented. Participants were provided with unlimited time to decide whether an item was old (previously studied) or new (not previously studied) by pressing either the "f" or "j" key. Again, each trial was followed by a blank screen for 250 ms. The key configuration was counterbalanced across participants within each group.

Results

Retrieval practice

In the noncompetitive retrieval practice task, participants retrieved, on average, 83.3 % of the category names (cued-recall group, *M* = 85.0 %, first block only = 83.8 %; recognition group, *M* = 81.6 %, first block only = 80.2 %; *t*s < 1).[1] In the competitive retrieval practice task, participants completed, on average, 74.1 % of the word stems (cued-recall group, *M* = 73.8 %, first block only = 71.1 %; recognition group, *M* = 74.4 %, first block only = 72.8 %; *t*s < 1).

Cued recall

In the cued-recall test, retrieval practice items (Rp+) were significantly better remembered, as compared with nonpracticed control items (Nrp+), in both the competitive condition, $t(39) = 5.50$, $p < .001$, $d = 0.87$, and the noncompetitive condition,

---

[1] Using a more liberal scoring method that did not require recalling the exact category name but also included truncations that had the same meaning as the correct answer resulted in an average retrieval practice performance of 91.3 % of category names (cued-recall group, M=91.8 %; recognition group, M=90.6 %; t<1).

$t(39) = 7.69$, $p < .001$, $d = 1.22$ (see Table 1). Furthermore, nonpracticed items of practiced categories (Rp−) were remembered worse, as compared with nonpracticed control items (Nrp−) in the competitive condition, $t(39) = -2.06$, $p < .05$, $d = 0.33$, and the noncompetitive condition, $t(39) = -2.40$, $p < .05$, $d = 0.39$. Retrieval-induced forgetting was therefore observed in both competitive and noncompetitive conditions.

Recognition

In the recognition test, false alarm rates for lures belonging to practiced categories (Rp lures) or nonpracticed categories (Nrp lures) were comparable in the noncompetitive condition (10.9 % vs. 9.9 %, $t < 1$) and in the competitive condition (8.8 % vs. 8.1 %, $t < 1$). As in previous research (e.g., Aslan & Bäuml, 2011; Gómez-Ariza et al., 2005; Verde & Perfect, 2011), d' was used as a measure of recognition accuracy for the four different item types.

The analyses showed a significant benefit for retrieval practice items (Rp+) recognition, as compared with nonpracticed control items (Nrp+) recognition, in the competitive condition ($d'_{Rp+} = 3.29$, $SE = 0.11$ vs. $d'_{Nrp+} = 2.75$, $SE = 0.11$), $t(71) = 4.64$, $p < .001$, $d = 0.55$, and in the noncompetitive condition ($d'_{Rp+} = 3.43$, $SE = 0.09$ vs. $d'_{Nrp+} = 2.61$, $SE = 0.11$), $t(71) = 8.60$, $p < .001$, $d = 1.01$. Critically, when the recognition performance of items that belonged to a practiced category but were not practiced during the retrieval practice phase (Rp−) were compared with nonpracticed control items (Nrp−), RIF was observed in the competitive condition ($d'_{Rp-} = 2.54$, $SE = 0.11$ vs. $d'_{Nrp-} = 2.77$, $SE = 0.12$), $t(71) = -2.16$, $p < .05$, $d = 0.25$, but not in the noncompetitive condition ($d'_{Rp-} = 2.62$, $SE = 0.13$ vs. $d'_{Nrp-} = 2.62$, $SE = 0.12$), $t(71) = -0.06$, $p = .955$.

Discussion

Replicating the standard RIF effect, competitive practice and cued recall resulted in RIF. However, since both theories, retrieval inhibition and strength-dependent competition, predicted forgetting, the source of forgetting is unclear. Noncompetitive practice and cued recall also resulted in RIF. Due to the absence of any retrieval competition during retrieval practice, only the strength-dependent competition account predicted forgetting. Using a modified procedure, this experiment replicates the findings of Raaijmakers and Jakab (2012) and others (Jonker & MacLeod, 2012).

As in previous studies (e.g., Aslan & Bäuml, 2010, 2011; Hicks & Starns, 2004), competitive practice and word recognition resulted in RIF. Assuming that recognition of an item does not require consideration of the associated category, only retrieval inhibition predicted forgetting. The strength-dependent competition account would have predicted forgetting if participants had attempted to reinstate learned cues. Finally, noncompetitive practice and word recognition did not result in

RIF. Retrieval inhibition predicted no forgetting due to the absence of retrieval competition during retrieval practice. Again, the strength-dependent competition account would have predicted forgetting only if participants had attempted to reinstate learned cues. Importantly, this null effect cannot be accounted for by low statistical power[2] or an ineffective manipulation (i.e., failure to strengthen the association between category and items), since the exact same manipulation caused forgetting in cued recall.

The results of Experiment 1 indicate that RIF caused by strength-dependent competition is not assessable by a recognition memory test. However, the above method deviated from Raaijmakers and Jakab's (2012) method in two potentially important aspects that may have influenced the strengthening of the word–category association. First, categories were mainly selected in terms of semantics rather than properties. Second, no feedback was given during retrieval practice. If a participant could not retrieve an Rp+ item the first time, the two consecutive practice trials may have also been ineffective in strengthening the word–category association. In additon, it might be possible to encourage participants to use the category of an item to determine whether the item has been learned or not. These issues were addressed in Experiment 2a.

Experiments 2a and 2b

Experiment 2a aimed to match the method closer to Raaijmakers and Jakab's (2012) method. To make the task more challenging and to focus learning on the category–word association, categories were selected only in terms of properties, rather than semantics. Since available norms did not meet this requirement or were not available in German, word materials were generated in a prestudy (see the Materials section). To ensure that participants were able to strengthen the category–word association during the retrieval practice phase, feedback on the correct answer was provided after each trial.

Finally, as was stated in the introduction, strength-dependent forgetting may be excluded only if participants do not try to reinstate learned cues and use these cues as retrieval aids during the recognition test phase (e.g., Camp et al., 2005, 2009; Verde & Perfect, 2011). A strong test of this assumption is to provide participants with learned cues and encourage them to use these cues for their old/new decision. In Experiment 2a, the category was shown for 2 s for each item, before the actual item was presented along with the category. This test is referred to as a *category-cued recognition* test. Experiment 2b was conducted to demonstrate RIF using the altered method of Experiment 2a and final cued-recall test.

---

[2] The power to detect a small-to-medium effect of d=0.035 was $1-\beta = .83$ (N=72, $\alpha=.05$).

In Experiment 2a, on the basis of the strength-dependent competition account, one would expect RIF if the participants use the category to make the old/new decision. Experiment 2b should show RIF if the altered method is effective in strengthening category–word associations.

Method

The method of Experiments 2a and 2b largely followed the method of Experiment 1. Therefore, only deviations are described.

Participants

In Experiment 2a, 72 university students (35 female) with a mean age of 22 years (range: 18–29 years) participated. In Experiment 2b, 40 university students (37 female) with a mean age of 21 years (range: 18–32 years) participated. In both experiments, participants took part in exchange for course credit, were native German speakers, and were tested in groups of 1–3.

Design

In both experiments, participants conducted a noncompetitive retrieval practice, and the *retrieval practice status* factor was manipulated within participants. In Experiment 2a, the final test included a category-cued recognition test, whereas the final test in Experiment 2b included a cued-recall test.

Materials

A total of 78 university students (52 female, age range, 18–30 years; mean age of 23 years) generated nouns for the following categories: *green*, *round*, *swim*, *soft*, *fly*, *wood*, *cold*, and *loud*. Participants were given 40 s to generate as many exemplars as possible per category. For each category, items were ranked on the basis of the number of participants who noted the item. The final materials for Experiment 2a consisted of six target items per category with low-to-medium taxonomic frequency (*Med* = 43), three frequent exemplars as frequent lures (*Med* = 19), and three exemplars as infrequent lures (*Med* = 72). The target items were between 4 and 13 letters long (*Med* = 6), and lure items were between 4 and 11 letters long (*Med* = 6). For Experiment 2b, eight target items were replaced to avoid two items in the same category beginning with the same initial letter. However, taxonomic frequency (*Med* = 43) and item length (range, 4–13 letters; *Med* = 6) were kept identical.[3] For both experiments, category sets and word sets were created and counterbalanced as in Experiment 1.

---

[3] All materials can be requested from the author.

Procedure

The study phase was as in Experiment 1, but each item was presented only for 2.5 s. The retrieval practice phase was as the noncompetitive practice in Experiment 1, but participants were provided with the correct answer for each category–word pair for 2 s (cf. Raaijmakers & Jakab, 2012).

The final memory test in Experiment 2a was identical to the recognition test in Experiment 1, with one exception. Participants were told that they would be shown the category of the next item for some time before being presented with the item and that this information was intended to assist them in the old/new decision. In the test, the category name was displayed for 2 s ("WOOD–") in between the fixation cross and the item. After 2 s, the item was presented along with the category ("WOOD–Cabinet"), and participants were provided with unlimited time to decide whether an item was old or new. The final memory test in Experiment 2b was identical to the cued-recall test in Experiment 1.

Results

Retrieval practice

In Experiment 2a and 2b, participants retrieved, on average, 96.6 % (first block only = 90.6 %) and 95.8 % (first block only = 89.2 %) of the category names, respectively. The means indicate that most errors were made in the first block.

Final memory test

In Experiment 2a, false alarm rates for lures belonging to practiced categories (Rp lures) or nonpracticed categories (Nrp lures) were comparable (5.6 % vs. 5.6 %, $t < 1$). Recognition of retrieval practice items ($d'_{Rp+} = 4.01$, $SE = 0.08$) was significantly better, as compared with nonpracticed control items ($d'_{Nrp+} = 2.70$, $SE = 0.10$), $t(71) = 12.64$, $p < .001$, $d = 1.49$ (see Table 1). However, recognition of nonpracticed items of practiced categories ($d'_{Rp-} = 2.68$, $SE = 0.10$) did not decrease, as compared with recognition of nonpracticed control items ($d'_{Nrp-} = 2.77$, $SE = 0.11$), $t(71) = -1.04$, $p = .30$.

In Experiment 2b, retrieval practice items (Rp+) were significantly better remembered than nonpracticed control items (Nrp+), $t(39) = 7.43$, $p < .001$, $d = 1.18$ (see Table 1). In addition, nonpracticed items of practiced categories (Rp−) were remembered worse than nonpracticed control items (Nrp−), $t(39) = -2.47$, $p < .05$, $d = 0.39$.

Discussion

The results of Experiment 2a show that RIF was absent if memory was tested using a category-cued recognition test. Again, this result is unlikely to be due to low statistical power.

Table 1 Results of Experiments 1, 2a, and 2b: Percentages of remembered items or "old" responses as a function of final test (cued recall, recognition), retrieval practice situation (competitive, noncompetitive), and retrieval practice status (Rp+, Nrp+, Rp−, Nrp−)

		Retrieval Practice Status					
		Rp+	Nrp+	Rp−	Nrp−	Rp Lure	Nrp Lure
Experiment 1							
Cued recall (% remembered items)	Competitive practice	67.8 (3.0)	42.7 (3.5)	44.0 (3.0)	51.5 (3.2)	–	–
	Noncompetitive practice	64.6 (2.3)	41.9 (2.4)	41.0 (2.9)	49.8 (2.8)	–	–
Recognition (% "old" responses)	Competitive practice	92.7 (1.1)	82.3 (2.0)	79.6 (1.9)	82.9 (1.8)	8.8 (1.0)	8.1 (0.9)
	Noncompetitive practice	95.9 (0.7)	81.9 (1.8)	82.2 (1.8)	82.5 (1.8)	10.1 (1.2)	9.9 (1.0)
Experiment 2a	Noncompetitive practice	97.7 (0.4)	77.7 (1.7)	74.7 (2.1)	77.4 (2.2)	5.6 (0.9)	5.6 (0.8)
Experiment 2b	Noncompetitive practice	61.7 (2.7)	27.5 (2.3)	34.4 (3.0)	42.9 (2.2)	–	–

Note. "Old" responses to Rp and Nrp items are hits; "old" responses to lure items are false alarms. Retrieval practice resulted in practiced items of practiced categories (Rp+) and their respective control items consisting of nonpracticed items of nonpracticed categories (Nrp+) and also resulted in nonpracticed items of practiced categories (Rp−) and their respective control items consisting of nonpracticed items of nonpracticed categories (Nrp−). The differences between Nrp+ and Nrp− items in cued-recall tests are due to controlled output order (Nrp− always before Nrp+ items). Standard errors are in parentheses.

In addition, Experiment 2b demonstrates that the modified method causes reliable RIF if a cued-recall test is applied as the final test. Overall, the results replicate the findings in relation to noncompetitive retrieval practice of Experiment 1, using a method that was very close to that in Raaijmakers and Jakab (2012) and a recognition test that encouraged participants to use the category to make an old/new decision.

General discussion

The results of the present study show reliable RIF using competitive retrieval practice and a cued-recall test or recognition test. For noncompetitive retrieval practice, RIF was observed in cued-recall tests, but not in an old/new recognition test (Experiment 1) or in an old/new-category–cue recognition test (Experiment 2a), despite reliable RIF occurring in the case of category-plus-initial-letter cued-recall tests.

In line with previous research, competitive retrieval practice and a cued-recall test (e.g., Anderson et al., 2000; Hanslmayr et al., 2010; Staudigl et al., 2010) or a recognition test (e.g., Aslan & Bäuml, 2010, 2011; Gómez-Ariza et al., 2005; Hicks & Starns, 2004; Spitzer & Bäuml, 2007) resulted in RIF. Similarly, noncompetitive retrieval practice, using a procedure that maximizes cue–item associations, and a cued-recall test resulted in RIF (Jonker & MacLeod, 2012; Raaijmakers & Jakab, 2012). The present results also correspond with Aslan and Bäuml's (2010) results, which show that individuals with limited inhibitory control (i.e., kindergarteners) do not show RIF in an old/new recognition test but do in a category-cued-recall test. Aslan and Bäuml (2010) made the argument that only the latter test involves interference; the recognition test is interference free. The present results support this argument, because no RIF was observed

in recognition tests despite noncompetitive retrieval practice that produces reliable RIF in cued-recall tests.

Theoretically, the present result pattern suggests that, in the standard retrieval practice paradigm, the source of forgetting cannot be attributed solely to retrieval inhibition, as proponents of such an account have argued (e.g., Anderson, 2003). An integration of both mechanisms has been mentioned by several authors (e.g., Aslan & Bäuml, 2010; Verde, 2009, 2012). One could assume that retrieval inhibition has its effect mainly during retrieval practice and that strength-dependent competition has its effect mainly during the final memory test. In the standard paradigm, during retrieval practice, the association of practiced words to their categories is strengthened, and interfering words are inhibited. During the final memory test, RIF is caused by (1) strong items interfering with the retrieval of the weaker target items, as well as (2) less active representations of these target items due to retrieval inhibition. Forgetting due to one of the two mechanisms is absent if either mechanism (1) is not supported during retrieval practice (i.e., no strengthening of category–word associations or no retrieval competition) or (2) is not supported during the final memory test (i.e., no retrieval competition). The present results support this idea, because RIF was observed in conditions in which at least one mechanism should have been effective yet was not observed in the condition involving noncompetitive practice and a recognition test. In this condition, the category retrieval practice supported the strengthening of category–word associations, but the recognition test did not support strength-dependent competition. Therefore, strengthened items appeared to not interfere during word recognition. This conclusion, however, is based on the assumption that recognition tests do not support retrieval competition.

As was mentioned in the introduction, results from other paradigms indicate that strength-dependent competition can, in

principle, affect recognition memory (Criss et al., 2011). Theoretically, the SAM-REM model for free recall (Malmberg & Shiffrin, 2005) can predict RIF in the present noncompetitive retrieval practice + recognition test conditions. However, in addition to the present results, Malmberg (2008) reported an experiment following the procedure of Malmberg and Shiffrin (2005) but did not observe a list strength effect in a single-item recognition test. Although the mechanisms of SAM-REM are more refined than in previous accounts (Rundus, 1973), the model relies, nonetheless, on the assumption of competition during the final test phase. In the present experiments, it may be questionable whether competition was present in the recognition test of Experiment 1. However, the category-cued recognition test of Experiment 2a should have caused competition, since the category-cues were present before the item and while the old/new decision was made. This suggests either that strengthened category–word associations did not cause interference during the recognition tests of these experiments or that participants successfully ignored the influence of strengthened-dependent interference. The present results support the idea that recognition tests are free from competition, but of course, this issue needs further research and a detailed analysis of the existing literature, which is beyond the scope of the present article.

Methodologically, the results of the present study suggest that recognition tests are a suitable way of distinguishing the contribution of retrieval inhibition and strength-dependent competition to RIF. Previous research using recognition tests only assumed that word recognition is an interference-free test and that RIF can be attributed solely to retrieval inhibition (e.g., Aslan & Bäuml, 2011; Hicks & Starns, 2004), although this assumption has been doubted (e.g., Verde & Perfect, 2011). Using the same noncompetitive retrieval practice procedure, RIF was observed for cued-recall tests but was absent in a standard recognition test (Experiment 1) and in a category-cued recognition test (Experiment 2a). On the other hand, a competitive retrieval practice produced RIF in both tests. On the basis of the present findings, it is recommended that one use old/new recognition tests instead of cued-recall tests as final memory tests when investigating inhibition-based RIF.

To summarize, the results of the present study support the idea that RIF can be caused by retrieval inhibition and strength-dependent competition. However, at least in the context of RIF, old/new recognition tests seem to be a reliable way of distinguishing between the different mechanisms as the cause of forgetting. Future research questions should attempt to separate the contribution of both mechanisms to RIF in the standard competitive retrieval practice.

References

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

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

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

Aslan, A., & Bäuml, K. H. (2011). Individual differences in working memory capacity predict retrieval-induced forgetting. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 37,* 264–269.

Aslan, A., & Bäuml, K. H. (2010). Retrieval-induced forgetting in young children. *Psychonomic Bulletin & Review, 17*(5), 704–709.

Bäuml, K. H., & Aslan, A. (2004). Part-list cuing as instructed retrieval inhibition. *Memory & Cognition, 32,* 610–617.

Buratto, L. G., & Lamberts, K. (2008). List strength effect without list length effect in recognition memory. *The Quarterly Journal of Experimental Psychology, 61*(2), 218–226.

Camp, G., Pecher, D., & Schmidt, H. (2005). Retrieval-induced forgetting in implicit memory tests: The role of test awareness. *Psychonomic Bulletin & Review, 12,* 490–494.

Camp, G., Pecher, D., Schmidt, H., & Zeelenberg, R. (2009). Are independent probes truly independent? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35,* 934–942.

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

Gómez-Ariza, C. J., Lechuga, M. T., Pelegrina, S., & Bajo, M. T. (2005). Retrieval-induced forgetting in recall and recognition of thematically related and unrelated sentences. *Memory & Cognition, 33,* 1431–1441.

Criss, A. H., Malmberg, K. J., & Shiffrin, R. M. (2011). Output interference in recognition memory. *Journal of Memory and Language, 64*(4), 316–326.

Hanslmayr, S., Staudigl, T., Aslan, A., & Bäuml, K. H. (2010). Theta oscillations predict the detrimental effects of memory retrieval. *Cognitive, Affective, & Behavioral Neuroscience, 10*(3), 329–338.

Hicks, J. L., & Starns, J. J. (2004). Retrieval-induced forgetting occurs in tests of item recognition. *Psychonomic Bulletin & Review, 11,* 125–130.

Jonker, T. R., & MacLeod, C. M. (2012). Retrieval-induced forgetting: Testing the competition assumption of inhibition theory. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale, 66*(3), 204–211.

Kinnell, A., & Dennis, S. (2011). The list length effect in recognition memory: An analysis of potential confounds. *Memory & Cognition, 39*(2), 348–363.

Malmberg, K. J. (2008). Recognition memory: A review of the critical findings and an integrated theory for relating them. *Cognitive Psychology, 57*(4), 335–384.

Malmberg, K. J., & Shiffrin, R. M. (2005). The "one-shot" hypothesis for context storage. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*(2), 322–336.

Mannhaupt, H.-R. (1983). Produktionsnormen für verbale Reaktionen zu 40 geläufigen Kategorien. *Sprache & Kognition, 4,* 264–278.

Mensink, G. J., & Raaijmakers, J. G. (1988). A model for interference and forgetting. *Psychological Review, 95,* 434–455.

Raaijmakers, J. G., & Jakab, E. (2012). Retrieval-induced forgetting without competition: Testing the retrieval specificity assumption of the inhibition theory. *Memory & Cognition, 40,* 19–27.

Raaijmakers, J. G. W., & Jakab, E. (2013). Rethinking inhibition theory: On the problematic status of the inhibition theory for forgetting. *Journal of Memory and Language, 68*(2), 98–122.

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

Rundus, D. (1973). Negative effects of using list items as recall cues. *Journal of Verbal Learning and Verbal Behavior, 12,* 43–50.

Spitzer, B., & Bäuml, K. (2007). Retrieval-induced forgetting in item recognition: Evidence for a reduction in general memory strength. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33,* 863–875.

Staudigl, T., Hanslmayr, S., & Bauml, K. (2010). Theta oscillations reflect the dynamics of interference in episodic memory retrieval. *Journal of Neuroscience, 30*(34), 11356–11362.

Storm, B., & Levy, B. (2012). A progress report on the inhibitory account of retrieval-induced forgetting. *Memory & Cognition, 40,* 827–843.

Van Overschelde, J. P., Rawson, K. A., & Dunlosky, J. (2004). Category norms: An updated and expanded version of the Battig and Montague (1969) norms. *Journal of Memory and Language, 50,* 289–335.

Veling, H., & van Knippenberg, A. (2004). Remembering can cause inhibition: Retrieval-induced inhibition as cue independent process. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30*(2), 315–318.

Verde, M. F. (2009). The list-strength effect in recall: Relative-strength competition and retrieval inhibition may both contribute to forgetting. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35*(1), 205–220.

Verde, M. F. (2012). Retrieval-induced forgetting and inhibition: A critical review. In B. H. Ross (Ed.), *The psychology of learning and motivation* (Vol. 56, pp. 47–80). New York: Academic Press.

Verde, M., & Perfect, T. (2011). Retrieval-induced forgetting in recognition is absent under time pressure. *Psychonomic Bulletin & Review, 11,* 1166–1171.