Journal of Experimental Psychology:
Learning, Memory, and Cognition
2009, Vol. 35, No. 1, 286–291

0278-7393/09/$12.00 DOI: 10.1037/a0014363

Retrieval-Induced Forgetting in a Category Recognition Task

Bernhard Spitzer and Karl-Heinz Bäuml
Regensburg University

Prior work on retrieval-induced forgetting showed that retrieving a subset of formerly studied items can impair item recognition of related, nonretrieved material. Here it was investigated whether retrieval practice can also impair the items' recognition as a member of a studied category. Subjects studied preexperimentally unrelated words that were categorized by their font colors, then practiced retrieving a subset of the words using a word stem completion task. Finally, an episodic category recognition test based on confidence ratings was applied. Receiver operating characteristic analysis of the data demonstrated a reliable impairment in subjects' recognition of the nonretrieved items' category relative to control items. The result indicates that retrieval-induced forgetting is not restricted to item recognition but generalizes to category recognition tasks. Inhibitory as well as noninhibitory explanations of the finding are discussed.

Keywords: episodic memory, retrieval-induced forgetting, recognition, category reduction

Retrieval-induced forgetting (RIF) refers to the finding that retrieving a subset of previously studied items can cause later forgetting of the nonretrieved material. Such forgetting has been demonstrated in both episodic and semantic memory tasks and for both verbal and visual material (for reviews, see Anderson, 2003; Bäuml, 2008). For instance, using verbal material, Anderson, Bjork, and Bjork (1994) found that when subjects study categorized material (e.g., *FRUIT–orange*, *FRUIT –apple*, *INSECT–bee*) and subsequently practiced retrieving half of the items from half of the studied categories (e.g., *FRUIT* or ___), later recall of the unpracticed material (*apple*) is impaired, relative to control items from the unpracticed categories (*bee*). Using visual material, Ciranni and Shimamura (1999) studied RIF in a purely episodic task. They used color as a grouping factor and asked subjects to learn the locations of differently colored objects. Retrieval practice on the locations of half of the objects from a color category improved later recall of the retrieved objects' locations but impaired recall of the nonretrieved objects' locations (for a demonstration of possible beneficial effects of retrieval practice on recall of nonretrieved material, see Chan, McDermott, & Roediger, 2006).

Detrimental effects of retrieval practice have been observed across a wide range of memory tests, including category-cued recall tasks (e.g., Anderson et al., 1994; Bäuml, Zellner, & Vilimek, 2005) and tests using so-called independent probes, that is, probes that were not used in previous phases of the experiment (e.g., Anderson & Bell, 2001; Saunders & MacLeod, 2006; but see Camp, Pecher, & Schmidt, 2007). In particular, RIF has repeatedly been demonstrated in tests assessing recognition memory (Gómez-Ariza, Lechuga, Pelegrina, & Bajo, 2005; Hicks & Starns, 2004; Spitzer & Bäuml, 2007; Spitzer, Hanslmayr, Opitz, Mecklinger, & Bäuml, in press; Starns & Hicks, 2004; Verde, 2004).

Inhibitory Versus Noninhibitory Accounts of RIF

RIF is often assumed to result from inhibitory control mechanisms that operate during retrieval practice (e.g., Anderson, 2003). Related not-to-be-retrieved items are thought to interfere and be inhibited to reduce interference during attempts to retrieve the target material. As a result of this inhibition, memory representations of the affected items are assumed to be reduced, leading to impaired performance on subsequent memory tests. The forgetting observed in independent-probe tests and recognition tests is consistent with this view. In particular, in a study using the remember–know procedure as well as an analysis of receiver operating characteristics (ROCs), evidence was reported that RIF in item recognition is due to a reduction in the general memory strength for unpracticed items (Spitzer & Bäuml, 2007).

RIF has also been explained by noninhibitory mechanisms. Noninhibitory accounts mostly explain the forgetting via mechanisms that act during the final memory test. Blocking theories, for instance, posit that practicing some members of a memory set results in a final recall situation where the strengthened items will come to mind more easily and block access to the weaker, unpracticed items that share the same cue (e.g., Williams & Zacks, 2001). Although blocking can explain the forgetting in category-cued recall tests, in which category information is provided and item-specific cues are absent, the account is harder to reconcile with the results from independent-probe and item recognition tests, in which no categorical information is provided (see Bäuml, 2008). The account also disagrees with the demonstration that RIF reflects a retrieval-specific effect (Anderson, Bjork, & Bjork, 2000; Bäuml, 2002; Ciranni & Shimamura, 1999).

Bernhard Spitzer and Karl-Heinz Bäuml, Department of Experimental Psychology, Regensburg University, Regensburg, Germany.

The research reported in this article was supported by Grant FOR 448 from the German Research Foundation (DFG) to Karl-Heinz Bäuml. We thank R. Schmidtner for her help in the experimental work and J. Chan for valuable comments on the article.

Correspondence concerning this article should be addressed to Karl-Heinz Bäuml, Department of Experimental Psychology, Regensburg University, 93040 Regensburg, Germany. E-mail: karl-heinz.baeuml@psychologie.uni-regensburg.de

From Item Memory to Category Memory

Prior work on RIF has focused on memory tests in which recall or recognition of target items themselves was required. These studies suggest that, for instance, if one had been introduced to a group of researchers from different universities (i.e., Morris and John from Berkeley and Philip and David from Oxford) and repeatedly retrieved one of the two Berkeley researchers' names (e.g., Morris), this might lead to failures to remember or recognize the name of the other Berkeley researcher (e.g., John). However, in real-life situations, people are not only asked to remember or recognize individual people or items but rather are also confronted with categorical questions, like "What is John's affiliation?" (category recall) or "Is John from Berkeley?" (category recognition). Whether prior retrieval of related items can impair memory for such categorical information in a similar way as it can impair memory for the target items themselves has not been addressed to date.

On the basis of the inhibitory account of RIF, impairments in memory for the unpracticed material's category membership may not be expected. Indeed, results from prior work suggest that features the unpracticed material shares with the practiced material are not subject to inhibition (Anderson, Green, & McCulloch, 2000; Bäuml & Hartinger, 2002; Bäuml & Kuhbandner, 2003). Because practiced and unpracticed material belong to a common category and thus share the same categorical features, retrieval practice should not affect the strength of the unpracticed material's categorical features and its category membership. As a result, no impairment in a category recognition task may arise.

In contrast, on the basis of the blocking account of RIF, impairments in memory for the unpracticed material's category membership may well be expected. Following this account, retrieval practice may lead to the blocking of unpracticed material at test. Although such blocking is quite unlikely to occur in item recognition tests, in which no categorical information is provided, it might play a role in categorical recognition tests, in which such information is a crucial part of the memory test. Indeed, the categorical information might activate the (relatively strong) practiced material, which might then block recognition of the (relatively weak) unpracticed material and thus reduce its category recognition (for more detailed explanations, see the Discussion section).

The Present Study

To the best of our knowledge, this study is the first one to examine whether the well-established detrimental effects of retrieval practice on recognition of unpracticed material itself (e.g., Hicks & Starns, 2004; Spitzer & Bäuml, 2007) extend to the material's category recognition. To address the issue, we report an experiment in which subjects studied preexperimentally unrelated words that were categorized according to their font color. After studying the material, subjects practiced retrieving half of the items from half of the color categories. After a distractor task, a category recognition test based on confidence ratings was applied. The studied items were presented in either their original or reversed font colors, and subjects had to indicate for each single item whether it was shown as member of its original color category. The rating data were analyzed using the ROC procedure (e.g., Wixted, 2007; Yonelinas, 2002). The results indicate whether the previous findings from item recognition tasks generalize to the present category recognition task.

Method

Subjects

Subjects were 72 young adults (19–34 years old) who were paid €6 (approximately $7.50) for participating in the experiment. They were tested individually in sessions that lasted approximately 25 min.

Material

The stimuli were 32 semantically unrelated German words chosen from different semantic categories (Battig & Montague, 1969; Mannhaupt, 1983). None of the items were preexperimentally associated with any particular color.

Design and Procedure

The experiment consisted of three main phases: a study phase, a retrieval practice phase, and a category recognition test phase. The 32 words were arbitrarily divided into four sets of eight items. For each subject, the font colors red, green, blue, and yellow were randomly assigned to the four sets of study items. From two of the four color categories, four items were used in the retrieval practice phase, resulting in three types of items: practiced items (P+ items), unpracticed items belonging to the same color categories as the practiced items (P− items), and control items from the unpracticed color categories (C items). Assignment of the items to retrieval practice was counterbalanced across subjects. All studied items were used in the category recognition test phase.

Subjects were tested individually in quiet surroundings, seated in front of a 17-in. TFT computer screen. They were instructed to memorize each item's color on the study list and were told that memory for the items' colors would be tested later, as all words would be shown again at test in potentially different colors. Each item was presented for 1,500 ms in its assigned font color, followed by a 500-ms blank screen. The order of the items was block randomized, that is, a random sequence of eight blocks consisting of one randomly selected exemplar from each of the four color categories was presented to the subjects with the constraint that no item in the sequence shared the next exemplar's color. Additionally, three buffer items were added at the beginning and at the end of the list, and the buffer items were also randomly assigned to one of the four font colors. After presentation of the study list, the study procedure was identically repeated with another random serial order of the study items. Directly after the study phase, subjects were instructed to count backward from 300 by threes for approximately 60 s.

Subsequently, the retrieval practice phase started. For each of the 8 (2×4) to-be-practiced items, the item's two initial letters were displayed in the same color as had been used in the study phase. Subjects were instructed to use the color information as a cue to complete the word stem with a studied item. The experimenter noted the subject's verbal response on a prepared data sheet, and the subject could proceed to the next item by pressing a key. The serial order of the word stems was quasi-random, with alternating font colors. For the following 5 min, a series of arithmetic problems was presented as a distractor task.

In the final category recognition test phase, half of the 32 studied items were displayed in the same font color in which they were studied (i.e., target, correct color), and the other half were displayed in one of the other three font colors (i.e., lure, incorrect color). The assignment of the items to correct or incorrect font colors at test was counterbalanced, and it was assured that for each item type (P+, P−, C), equally many items served as targets and lures. Each test item was presented in the middle of the screen, together with a schematically depicted 6-point rating scale ranging from 1 = *definitely correct color* to 6 = *definitely wrong color* in the lower part of the screen. Subjects were instructed to enter their responses via the digits of the PC keyboard. As soon as any numerical response was entered, the next test item was presented on the screen. The serial order of the test items was again block randomized, with the additional constraint that no type of item requiring the same correct response (target or lure) appeared more than four times in a row.

Results

Retrieval Practice Phase

Subjects, on average, successfully completed 59.5% (*SE* = .02) of the word stems presented during retrieval practice.

Recognition Test

Descriptive analysis of recognition performance. For each of the three item types (P+, C, P−), overall hit and false alarm rates were determined by collapsing subjects' responses across the three correct color levels of the rating scale (i.e., 1 = *definitely correct color*; 3 = *probably correct color*). The mean hit rate was .92 (*SE* = .02) for practiced material, .85 (*SE* = .02) for control material, and .81 (*SE* = .02) for unpracticed material. The mean false alarm rate was .24 (*SE* = .03) for practiced material, .22 (*SE* = .02) for control material, and .27 (*SE* = .03) for unpracticed material. Recognition accuracy as indexed by hit rate minus false alarm rate thus was .68 (*SE* = .03) for practiced material, .63 (*SE* = .03) for control material, and .54 (*SE* = .03) for unpracticed material.

As is evident from the descriptive data, subjects generally achieved high performance levels in the category recognition task, indicating that an episodic categorization of the study material had been successfully established. Moreover, the category recognition data exhibit a pattern of RIF with increased recognition performance for practiced material and decreased performance for unpracticed material, relative to control material. For a detailed statistical analysis of these effects, we used a signal detection approach that allowed us to objectively quantify memory performance while controlling for possible confounds of response bias.

Signal detection analysis. For signal detection analysis of the recognition data, hit rates and false alarm rates were cumulated over the five criterion points of the confidence rating scale, starting with the most confident criterion point (i.e., 1 = *definitely correct color*). The ROCs obtained by plotting the cumulative hit rates against false alarm rates for each of the three item types are illustrated in Figure 1A and Figure 1B. For statistical analysis, we fitted a signal detection model (equal variance) to the ROC data, assuming that recognition in this experiment was based on a single source of memorial information (i.e., [general] memory strength). Following detection theory, subjects should respond with a given level of confidence whenever their assessment of the memory strength of an item's membership to a certain color category exceeded the response criterion, c, associated with that confidence level (Figure 1C). The target items' memory strength is then given by the distance between the means of the underlying strength distributions for targets (correct color) and lures (incorrect color; d'). Assuming equal variances of the underlying strength distributions, the model has six free parameters (d' and Response Criteria c_1–c_5) to fit the 10 data points in a six-alternative forced choice rating experiment (hits and false alarms for Confidence Levels 1–5). Thus, the model has four degrees of freedom for testing its goodness of fit. The model parameters were estimated using maximum likelihood techniques, which also allow for statistical testing (for technical details, see the Appendix of Spitzer & Bäuml, 2007).

The equal-variance signal detection model's fit to the data is graphically illustrated in Figure 1A and Figure 1B, and the parameter estimates and goodness-of-fit statistics are given in Table 1. The model described the data of all three item types well, all $\chi^2(4)$s < 3.40, all ps $> .50$.[1] The estimates for d' exhibited the standard pattern of RIF: P+ (2.14) > C (1.84) > P− (1.57; see Figure 1D). Likelihood ratio tests confirmed that, relative to control items, memory strength as measured by d' was significantly higher for practiced items, $\chi^2(1, N = 864) = 14.43$, $p < .001$, and reliably lower for unpracticed items, $\chi^2(1, N = 864) = 15.94$, $p < .001$.

As is evident from the ROCs shown in Figure 1A and Figure 1B, the response criteria for practiced material were relatively liberal (i.e., shifted to the right) compared with the response criteria for control and unpracticed material. Statistical analysis of the model parameters confirmed that the estimated response criteria c_1–c_4 were significantly smaller (i.e., more liberal) for practiced material compared with control material, all $\chi^2(1)$s > 6.20, all ps $< .05$. In contrast, the response criteria for unpracticed material were about the same as for control material, all $\chi^2(1)$s < 1. The analysis indicates that compared with control and unpracticed material, subjects showed a greater tendency to rate practiced material as old, regardless of whether the items' color configuration was actually studied.

Fitting more general signal detection models. The equal-variance signal detection model implies that the strength distributions underlying targets and lures are normal and have equal variances (i.e., $\sigma = V_{\text{targets}}/V_{\text{lures}} = 1$; see Figure 1C). As a consequence, z-ROCs predicted by the model are linear (through normality) and have a slope of $1/\sigma = 1$ (through homogeneity of variances; see Figure 1B). For a statistical test of the validity of the equal-variance assumption, we fitted the more general unequal-variance signal detection model to the data, allowing parameter σ to vary freely. Across conditions, the estimates for parameter σ were not

[1] The retrieval practice paradigm yields relatively few observations per subject and item type. Thus, fitting the data on an individual-subject level would cause major distortions of the model's parameter estimates. Accordingly, the modeling analysis in the present study was performed on aggregate data (see Spitzer & Bäuml, 2007).

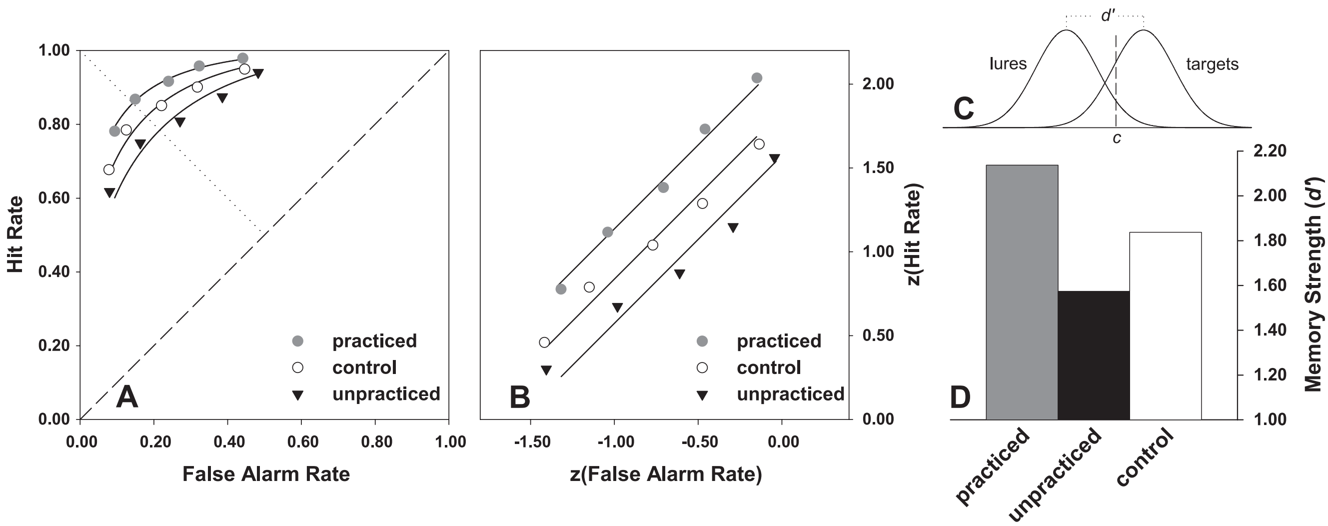

Figure 1. A: Category recognition receiver operating characteristics (ROCs) depicting the cumulative hit and false alarm rates for the three item types (practiced, control, unpracticed). Recognition accuracy (hit rate − false alarm rate) is highest for practiced material and lowest for unpracticed material. Solid lines indicate theoretical ROCs predicted by the equal-variance signal detection model. B: *Z* transformation (*z*-ROCs) of the plot shown in Figure 1A. The equal-variance signal detection model predicts linear *z*-ROCs with a slope of 1 and an intercept of d'. C: Solid lines represent schematically depicted strength distributions for targets and lures as assumed by the equal-variance signal detection model. The dashed line represents the location of response criterion, *c*. According to the model, subjects respond with a certain level of confidence whenever their assessment of memory strength exceeds the response criterion associated with that confidence level. The dotted lines represent the distance between the two distributions' means (d'), indicating the target material's relative memory strength. D: Parameter estimates for memory strength (d') derived from the equal-variance signal detection model when fitted to the ROCs of the three item types. Relative to control material, memory strength is significantly increased for practiced material ($p < .001$) and significantly reduced for unpracticed material ($p < .001$).

significantly different from 1, all $\chi^2(1)$s < 1.70, all *p*s $> .15$, indicating that the category recognition ROCs conform well with the standard signal detection model's assumption of equal variances.

We also fitted the dual-process signal detection model to the data, assuming a thresholdlike recollection process in addition to a strengthlike signal detection familiarity process (Yonelinas, 2002). Technically, inclusion of a recollection parameter allows for detecting systematic nonlinearities in the *z*-ROC's high-confidence region. However, applied to the present data, the estimates for the recollection parameter were not significantly different from zero, all $\chi^2(1)$s < 1.80, all *p*s $> .15$. Thus, statistical model comparisons indicate that the data are well described by the equal-variance signal detection model, assuming that category recognition performance was based on a single source of memorial information (for technical details, see Spitzer & Bäuml, 2007).[2]

Guessing bias analysis. Signal detection analysis revealed reduced memory performance but no changes in response strategy or guessing for unpracticed material compared with control material. Still, the possibility remains that the observed difference in d' was due to biased guessing about the material's original color category. Indeed, if prior retrieval practice on a subset of the color categories had led subjects to prefer unpracticed colors when guessing about a nonretrieved item's original color category, such biased guessing could have artificially reduced discriminability for unpracticed compared with control material. In particular, a guessing preference for unpracticed colors would have artificially increased hits for control targets (always shown in unpracticed colors) and decreased hits for unpracticed targets (always shown in practiced colors). In addition, such biased guessing would have decreased false alarms for control lures (shown in practiced colors in the majority of cases) and increased false alarms for unpracticed lures (shown in practiced colors in the majority of cases).[3] A guessing preference for unpracticed colors would thus have artificially created a pattern of RIF, with increased d' for control material and decreased d' for unpracticed material.

[2] When the restriction that the recollection parameter is zero is imposed, the dual-process model is formally equivalent to an equal-variance signal detection model.

[3] Note that, at test, lures were created by randomly assigning half of the items of each item type (P+, C, P−) to any color category other than their original one. For example, if a control item was studied in a red font, it could be turned into a lure item by presenting it either in the wrong unpracticed color (e.g., blue) or in one of the practiced color categories (green or yellow, in the present example). Therefore, on average, control lures were in a third of cases presented in unpracticed colors and in two thirds of cases in practiced colors; in contrast, unpracticed lures were in two thirds of cases presented in unpracticed colors and in one third of cases in practiced colors.

Table 1
Parameter Estimates and Goodness-of-Fit of the Equal-Variance Signal Detection Model When Fitted to the ROCs of the Three Item Types

Item type	d'	c_1	c_2	c_3	c_4	c_5	$G^2(4)$
Practiced	2.14*	0.28*	−0.04*	−0.34*	−0.62*	−0.93	0.44
Control	1.84	0.47	0.17	−0.14	−0.42	−0.77	2.58
Unpracticed	1.57*	0.53	0.15	−0.14	−0.45	−0.75	3.34

Note. Asterisks indicate significant deviations from control performance.
* $p < .05$.

We examined this type of guessing bias by comparing the false alarm rates for lure pairings with practiced colors with the false alarms rates for lure pairings with unpracticed colors. For control lures, false alarm rates did not differ between practiced and unpracticed colors for any of the response criteria (all $ps > .25$). For unpracticed lures, false alarm rates were higher for practiced than unpracticed colors at Criterion Point 3 (i.e., the subjective old criterion; 0.29 vs. 0.20, $p < .05$) but did not differ for the other response criteria (all $ps > .15$). The observed decrement in memory performance for unpracticed material thus cannot be attributed to biased guessing. If anything, the memory strength for unpracticed material d' was even overestimated due to subjects' tendency to guess a practiced category when presented with unpracticed material.

Discussion

Prior work on RIF focused on memory tests in which subjects' memory of the target items themselves was assessed. Unpracticed items were forgotten over a wide range of memory tests, including tests assessing item recognition (e.g., Hicks & Starns, 2004; Spitzer & Bäuml, 2007). The results of the present study extend this prior work by demonstrating that the detrimental effect of retrieval practice is not restricted to item memory tests but generalizes to categorical memory tests, in which the task is to recognize a target as member of a specific episodic category. In fact, as indicated from an ROC analysis of the data, after retrieval practice, memory for an item's membership to a specific category can be poorer for unpracticed items than for control material.

Relation to Prior Work

Recognition of episodic material is sometimes assumed to rely on two memory codes, recollection and familiarity (e.g., Yonelinas, 2002), and sometimes to be entirely based on a single source of memorial information, the item's general memory strength (e.g., Wixted, 2007). In recent work, we demonstrated that RIF in item recognition is well described by a single-process view, indicating that retrieval practice reduces the general memory strength for unpracticed items (Spitzer & Bäuml, 2007). The present results generalize this finding and indicate that RIF in category recognition is well described by a single-process view as well. Thus, as holds true in item recognition, no dual-process view appears necessary to describe the forgetting in category recognition.

To some extent, the present results appear to disagree with previous work reporting detrimental effects of retrieval practice on recognition of paired associates (Verde, 2004). Using the remember–know procedure, Verde's study showed that retrieval practice affects subjects' recognition of unpracticed material at high levels of response confidence (remembering) but does not impair or even facilitate recognition at moderate levels of confidence (knowing). In contrast, in the present experiment, the recognition decrement was evident across all confidence levels of the ROC, indicating a more general impairment than was reported in the previous associative recognition work. This discrepancy in the findings may be due to qualitative differences between the present category recognition task and standard associative recognition tasks: Whereas ROCs from associative recognition experiments are typically strongly asymmetrical or even linear in shape (e.g., Kelley & Wixted, 2001; Yonelinas, 2002), the present category recognition ROCs are curvilinear and fairly symmetrical in shape, suggesting that the two types of test reflect different aspects of memorial information.

Like we did in the present study, Ciranni and Shimamura (1999) studied RIF using color as a grouping factor. Using colored objects and their locations as stimuli, they demonstrated that retrieval practice on half of the locations of the objects from a color category can impair later recall of the nonretrieved objects' locations. The present finding extends this prior result by showing that retrieval of a subset of perceptually grouped items can also impair recognition of unpracticed items' episodic category membership. Together, the two lines of research indicate that retrieval practice can lead to fairly broad memory deficits for episodically related unpracticed material.

Implications for Theoretical Accounts of RIF

The question arises of whether the present results are consistent with inhibitory or noninhibitory accounts of RIF. If inhibition is supposed to mediate not only the detrimental effect of retrieval practice on item recognition but also the observed impairment in category recognition, then inhibition must be assumed to weaken the unpracticed material's categorical features and its strength of category membership. However, prior work indicates that features that practiced and unpracticed items share, like categorical features, are not subject to inhibition (Anderson, Green, & McCulloch, 2000; Bäuml & Hartinger, 2002; Bäuml & Kuhbandner, 2003). The strength of the unpracticed material's categorical features, therefore, should not have been weakened in the present experiment, so, according to inhibition, category recognition should not have been impaired. It thus appears unlikely that the observed impairment was mediated by inhibition.

According to the blocking account of RIF, at test, the practiced items come to mind more easily and may block access to the weaker, unpracticed items that share the same cue (Williams & Zacks, 2001). Such blocking might have occurred in the present experimental situation, in which categorical information was a crucial part of the recognition test. The categorical information may have activated the practiced material, thus blocking recognition of the unpracticed material and leading to the observed reduction in category recognition. As a variant of this account, retrieval practice may also have increased the typicality of the practiced material within its category, which may have led to a decrease in the unpracticed material's relative typicality, thus impairing its category recognition.

The present experiment was designed to examine whether RIF generalizes to categorical memory tests but was not designed to ultimately distinguish between inhibitory and noninhibitory accounts of the forgetting. Still, the results favor the view that blocking rather than inhibition caused the RIF effect in the present category recognition task. If so, the results of this study suggest that different mechanisms mediate RIF effects in different experimental tasks: Whereas inhibition seems to mediate forgetting in item memory tasks (e.g., Anderson, 2003), blocking may mediate forgetting in category memory tasks. Prior work has already assigned both inhibition and blocking roles in RIF (e.g., Camp et al., 2007; Verde, 2009), particularly when category information is provided and item-specific cues are absent, as, for instance, occurs in category-cued recall tests (see Bäuml, 2008). This prior work, however, focused on item memory tasks. The present study thus goes beyond this prior work by suggesting that different mechanisms mediate RIF in item memory versus category memory tasks. Such a scenario offers a new, advanced view on explaining the detrimental effects of retrieval practice.

Conclusions

Over a wide range of memory tasks, previous work demonstrated detrimental effects of retrieval practice on memory for unpracticed material, indicating that recognition of a putative colleague's name John may be impaired if one previously retrieved other Berkeley researchers' names, like Morris. The results of the present study suggest that there may be another memory problem that comes along with retrieval practice. One's retrieval of the name Morris may impair categorical memories of John, as addressed in questions like, "What is John's affiliation?" or "Is John from Berkeley?" Item recognition and category recognition differ in experimental task and may assess different aspects of an item's complex representation in memory. Despite the differences, prior retrieval practice on related material can impair memory performance in both types of tasks.

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., 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.

Bäuml, K.-H. (2002). Semantic generation can cause episodic forgetting. *Psychological Science, 13,* 356–360.

Bäuml, K.-H. (2008). Inhibitory processes. In J. H. Byrne (Ed.-in-Chief) & H. L. Roediger III (Vol. Ed.), *Learning and memory: A comprehensive reference: Vol. 2. Cognitive psychology of memory* (pp. 195–220). Oxford, England: Elsevier.

Bäuml, K.-H., & Hartinger, A. (2002). On the role of item similarity in retrieval-induced forgetting. *Memory, 10,* 215–224.

Bäuml, K.-H., & Kuhbandner, C. (2003). Retrieval-induced forgetting and part-list cuing in associatively structured lists. *Memory & Cognition, 31,* 1188–1197.

Bäuml, K.-H., Zellner, M., & Vilimek, R. (2005). When remembering causes forgetting: Retrieval-induced forgetting as recovery failure. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31,* 1221–1234.

Battig, W. F., & Montague, W. E. (1969). Category norms of verbal items in 56 categories: A replication and extension of the Connecticut category norms. *Journal of Experimental Psychology, 80,* 1–64.

Camp, G., Pecher, D., & Schmidt, H. G. (2007). No retrieval-induced forgetting using item-specific independent cues: Evidence against a general inhibitory account. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33,* 950–958.

Chan, J. C. K., McDermott, K. B., & Roediger, H. L., III (2006). Retrieval-induced facilitation: Initially nontested material can benefit from prior testing of related material. *Journal of Experimental Psychology: General, 135,* 553–571.

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.

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

Kelley, R., & Wixted, J. T. (2001). On the nature of associative information in recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27,* 701–722.

Mannhaupt, H.-R. (1983). Reproduktionsnormen für verbale Reaktionen zu 40 geläufigen Kategorien [Production norms for verbal reactions to 40 familiar categories]. *Sprache und Kognition, 4,* 264–278.

Saunders, J., & MacLeod, M. D. (2006). Can inhibition resolve retrieval competition through the control of spreading activation? *Memory & Cognition, 34,* 307–322.

Spitzer, B., & Bäuml, K.-H. (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.

Spitzer, B., Hanslmayr, S., Opitz, B., Mecklinger, A., & Bäuml, K.-H. (in press). Oscillatory correlates of retrieval-induced forgetting in recognition memory. *Journal of Cognitive Neuroscience.*

Starns, J. J., & Hicks, J. L. (2004). Episodic generation can cause semantic forgetting: Retrieval-induced forgetting of false memories. *Memory & Cognition, 32,* 602–609.

Verde, M. F. (2004). The retrieval practice effect in associative recognition. *Memory & Cognition, 32,* 1265–1272.

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,* 205–220.

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

Wixted, J. T. (2007). Dual-process theory and signal-detection theory of recognition memory. *Psychological Review, 114,* 152–176.

Yonelinas, A. P. (2002). The nature of recollection and familiarity: A review of 30 years of research. *Journal of Memory and Language, 46,* 441–517.

Received March 18, 2008
Revision received September 24, 2008
Accepted September 25, 2008 ■