Mem Cogn (2013) 41:511–518
DOI 10.3758/s13421-012-0276-3

Interference resolution in retrieval-induced forgetting: Behavioral evidence for a nonmonotonic relationship between interference and forgetting

Attila Keresztes · Mihály Racsmány

Published online: 21 December 2012

Abstract Retrieving memories renders related memories less accessible. This phenomenon, termed *retrieval-induced forgetting* (RIF), is thought to be the result of processes that resolve interference during competitive retrieval. In several studies, researchers have manipulated the level of interference to test different theoretical accounts of RIF (e.g., inhibitory vs. non-inhibitory). However, the nature of how interference and RIF are related has not been systematically investigated. Here, we introduce a design that allows for assessing interference during competitive retrieval by measuring the recall RTs associated with target recall. Using such a design, we found that RIF occurred only when interference during competitive retrieval reached moderate levels, but not when it was too low or too high. This finding might indicate that low levels of interference do not trigger interference resolution, whereas interference resolution might fail when the interference reaches extremely high levels.

Keywords Retrieval · Forgetting · Retrieval-induced forgetting · Interference · Inhibition

Interference as a cause of forgetting has long captured the interest of scholars of memory. One specific question that has resurfaced in scientific discussions has concerned the way that interference during memory retrieval is resolved (for a review of interference theories from this perspective, see, e.g., Anderson & Bjork, 1994). The focus in these discussions was not solely on how interference causes memory failures during retrieval. Rather, it centered around the consequences of interference resolution. What happens to memory representations when a target memory is to be retrieved in the face of competing memories? How do we achieve retrieval of the correct target memory, and what happens to competing memories?

A. Keresztes (✉) · M. Racsmány
Department of Cognitive Science, Budapest University of Technology and Economics, Budapest, Hungary
e-mail: akeresztes@cogsci.bme.hu

In a seminal study, Anderson (2003) suggested an executive process—analogous to response override—that resolves interference by weakening memory representations that interfere with target memories at the time of recall. This weakened representation would be evident in the decreased probability of recall of the interfering memory when it is tested at a later time. This model was the first to hypothesize an active executive process that can act to weaken memory representations so that those memories become less accessible for retrieval.

Early on, Anderson, Bjork, and Bjork (1994) devised a procedure, the *retrieval practice paradigm*, that can separate a recall phase at time *t*, when interference from competing memories has to be resolved, from a recall phase at time *t*+1, when the accessibility of these competing memory representations can be measured. In this paradigm, participants are shown several category–member word pairs (e.g., *animal–tiger*, *furniture–couch*, and *animal–chicken*) and then practice retrieval of half of the members from half of the categories with category-plus-stem cues (e.g., *animal–ti…*?). Anderson et al. (1994) reasoned that, during retrieval practice, nonpracticed members of practiced categories (e.g., *chicken*) would interfere with the recall of practiced members, and therefore have to be inhibited. This should be evident from later testing, and that was exactly what they found: Participants' recall of nonpracticed members of practiced categories was worse than their recall of members of categories that had not appeared in the retrieval practice phase. Anderson et al. (1994) termed this effect *retrieval-induced forgetting* (RIF).

Inhibition, in this theoretical approach, is a process that operates when a relatively strong competing item interferes with the retrieval of a target memory. This approach involves three testable properties of RIF that are relevant for our study. First, RIF is interference dependent; that is, only items interfering with the retrieval of a target memory would suffer inhibition. Second, RIF is retrieval specific; that is, manipulating target strength without retrieval of the target would not induce competition-based forgetting. Third, RIF is strength independent; that is, even when targets are retrieved, target strength does not influence RIF (e.g., Anderson, 2003). Together, these assumptions imply that RIF is the product of executive processes that resolve interference during competitive retrieval.

The relationship between interference and interference resolution

Although the dependence of RIF on interference seems to be well established, we know little about how interference-resolving processes operate in the face of increasing interference. Proponents of different inhibitory accounts of RIF have conceived this relationship quite differently. Anderson (2003, p. 421) suggested that “The more strongly associated to the category an unpracticed competitor was, the more impairment was found.” This implies a linear function between interference and the result of inhibition: The more that an item interferes with retrieval, the more inhibition it suffers.

Bäuml, Pastötter, and Hanslmayr (2010, p. 1049) suggested that “very low levels of interference may not trigger inhibitory processes when competing material is retrieval practiced,” but otherwise the strength of the interference that a competing item causes during retrieval plays only a minor role in defining the level of forgetting. According to their interpretation, inhibition would be a process that kicks in only when interference reaches a certain threshold. They also theorized that, over this threshold level, the effect of inhibition would not change significantly with increasing interference.

Norman, Newman, and Detre (2007) programmed a neural network model of RIF in which increasing competitor strength increased the effect of inhibition, but over a certain point, this could lead to a decrease in its success (i.e., a decrease in RIF). Similarly, Anderson and Levy (2010) suggested that a positive linear relationship exists between the level of interference and inhibition demands, and a negative linear relationship between the level of interference and inhibition success. Together, these opposing relationships lead to a demand–success trade-off in which very low levels of interference do not lead to RIF, because the level of inhibition demand remains low, whereas interference can reach a level over which inhibition cannot be effective, resulting in above-baseline facilitation of competitors. The carryover assumption put forward by these authors states that RIF should be seen only for items that induce moderate levels of interference.

The lack of knowledge about the function relating interference to forgetting makes it hard to design tests that try to tap properties of RIF. Take a study that tries to provide evidence for interference dependence by including a group of items with strong taxonomic frequencies (supposedly inducing great interference) and another group with low taxonomic frequencies (supposedly inducing little interference; e.g., Anderson et al., 1994; Williams & Zacks, 2001). This study would reliably provide significant RIF differences between these two groups if the relationship between interference and forgetting was a simple linear one, as suggested by Anderson (2003). In the case of a threshold-like interference resolution process, as suggested by Bäuml et al. (2010), differences would only be found if the low-taxonomic-frequency words did not achieve a certain threshold at which inhibition kicks in. Moreover, if inhibition causes forgetting to decrease over a certain level of interference, as predicted by Norman et al. (2007) and Anderson and Levy (2010), one might see no differences between the two groups, because one of them could cause no interference at all, while the other one could cause too much interference. Studies using factorial designs might obtain contradictory results (see, e.g., Anderson et al. [1994] vs. Williams & Zacks [2001]) simply because the groups of words chosen to cause great or little interference are chosen on an arbitrary basis and without any knowledge of the underlying relationship between interference and the effect of interference resolution.

Another advantage of understanding how the effect of interference resolution changes as interference increases would be to design tests that are more sensitive to detect RIF. Such tests could focus only on memories that truly caused interference during memory retrieval, and thus that would be expected to suffer the results of interference resolution. Such sensitive tests would be very useful in settling some hot debates about the nature of interference-resolving processes in memory—for instance, to clarify whether RIF generalizes to novel, independent cues (for positive evidence, see, e.g., Aslan, Bäuml, & Pastötter, 2007; Levy, McVeigh, Marful, & Anderson, 2007; Saunders & MacLeod, 2006; for negative evidence, see, e.g., Camp, Pecher, & Schmidt, 2007; Perfect, Stark, Tree, Moulin, Ahmed and Hutter 2004).

Item-by-item RIF

Our goal in this study was to develop a test that could give an indication of how RIF changes as a function of competition during retrieval. Therefore, we needed a design that could provide data on how retrieval of each memory item was affected by interference during the retrieval practice phase. We set two objectives to achieve this goal. First, the design should be such that each item had an individual competitor that interfered with it. Second, we needed to collect data that at least indirectly would inform us about the amount of interference that a memory item suffers during its retrieval in the retrieval practice phase. For this second purpose, we chose to measure the reaction times (RTs) of target memory retrieval during the retrieval practice phase.

RTs have been used to measure the levels of interference caused by competing representations or processes in a number of paradigms—among others, negative priming (Tipper, 1985), repetition priming (Rajaram, Srinivas, & Travers, 2001), and the stop signal RT task (Logan & Cowan, 1984). Blaxton and Neely (1983) showed that RTs to generate the target exemplar were faster if the participant had first read other exemplars from the same category rather than exemplars from a different category. However, RTs were slower if the participants had first generated other exemplars from the same category.

RT data have rarely been collected in RIF studies. Anderson (2003, p. 439) suggested that "when the measure of interference is reaction time, the presence of multiple competitors or a single strong competitor should slow the recall of a target." Indeed, RTs have been used in RIF studies to measure the magnitude of the RIF effect (e.g., Racsmány & Conway, 2006; Veling & van Knippenberg, 2004; Verde & Perfect, 2011).

In a similar vein, RTs have been used to measure interference during retrieval practice. In one study (Levy et al., 2007, Exp. 2), the participants were split into two groups according to the interference that a memory suffered during retrieval practice. In this study, participants had to name pictures in their second language and were tested later using the same pictures in their first language. Levy et al. performed a median split of their sample based on the overall RT differences between the participants' performance in the first and second languages. The authors suggested that slower naming performance in the second than in the first language indicates poorer knowledge of the second language. On this basis, they hypothesized that participants with larger RT differences would need to resolve greater interference from the first language when naming pictures in their second language than would participants who have better knowledge of their second language. This would lead to greater RIF among poorer speakers than in the other group, and this is exactly what was found.

Kuhl, Dudukovic, Kahn, and Wagner (2007), measured RTs and activation in prefrontal areas during retrieval practice and correlated the amount of RIF with the decrease of these measures from the first to the third practice cycle. They found that the decreases in prefrontal activation, but not RTs, correlated positively with forgetting of the interfering memories. It is important to note that such a reduction is more a measure of successful interference resolution than of interference per se.

Here, we used a variation of the retrieval practice paradigm introduced by Anderson et al. (1994), in which only two items share the same category cue (and compete for retrieval) in every category. We did not manipulate interference in a factorial design, but rather used the retrieval practice RTs as an independent variable to assess the magnitude of interference. Of course, we do not assume that retrieval RTs only reflect interference. They are influenced by several other factors as well, such as target strength and number of competitors. In the Method section, we will discuss how we tried to control the variability of these potential factors.

Using such item-by-item RIF, we intended to reproduce findings supporting the interference dependence of RIF and to better understand how interference and the forgetting effect caused by interference resolution are related.

Method

Participants

A group of 64 students (age: $M = 21.81$ years, $SD = 2.12$; 32 women, 32 men) participated in the experiment for credit in partial fulfillment of an introductory psychology course requirement at Budapest University of Technology and Economics. The participants were tested individually in a quiet room in sessions that lasted for a maximum of 30 min. Due to a computer error, one participant could not complete the test phase. This participant's data were excluded from the analyses.

Materials

We used 22 categories with two members in each category, making a list of 44 word pairs. To induce the competitive retrieval that is supposed to be necessary to produce RIF, and to avoid moderation of the RIF effect (see Anderson, 2003), we followed strict selection criteria. To produce any RIF effect, it would be essential to have items in a category that would interfere with each other. Integration has been shown to counteract the RIF effect robustly (Anderson & McCulloch, 1999), and reducing the number of elements in a category increases the chances of integration (e.g., Camp et al., 2007). Since we used only two members per category, we had to take care to reduce the chances of integration.

Frequency and association data were drawn from the open-source *Frequency Dictionary of the Hungarian Webcorpus*, developed by BME Média és Szociológia Tanszék–Média Oktatási és Kutató Központ (Media Research Centre at the Department of Sociology and Communications of Budapest University of Technology and Economics; BME-MOKK, 2003). For a full description of the database, see Halácsy et al. (2004) and Kornai et al. (2006). We included categories that were not associated with each other (either semantically or phonetically) and that were themselves of moderate frequency. The category labels and targets were neutral words. Category members were moderate-frequency words, and within their category they had a moderate to high relative frequency. Category members that were either too typical or too rare were excluded. No member from a given category was associated with another member in another category, nor was it associated to another category cue. We made an effort to choose the two members of one category from different subcategories. To avoid cues that would uniquely refer to one target in semantic memory during the test phase, the first letter of each target was shared with at least one other low- or moderate-frequency category member that did not appear in the experiment. In contrast, to avoid extraexperimental interference during retrieval practice, we excluded words whose first two letters could be completed to create another category member not seen in the experiment. The first two letters had a moderate versatility; that is, a moderate number of words could be generated from the same two letters from semantic memory. We made an effort to reduce the number of words in which the first two letters made up or contained a syllable of the word.

After filtering possible materials through these selection criteria, we had a list of 88 words, including four words belonging to each of the 22 categories. In order to reduce item-based confounds in the RT data, we wanted to create a final list that would produce the least variation in baseline retrieval RTs. To this end, we ran a pilot study in which participants learned all 88 category–member pairs and then performed retrieval practice on all of the items once. To obtain the final list to be used in our experiment, we excluded two items per category on the basis of the retrieval practice results in this study. Using recall RTs, we excluded words that produced RTs that either were more than one standard deviation away from the group mean or differed substantially (more than 1,000 ms) from the group mean RT of their category. Using recall accuracy, we excluded both words that were recalled by every participant in the pilot and words that were recalled by less than 33 percent of the participants (around the lower and upper deciles of the 88 words; see the Table 2 for the final list of word pairs selected.)

We used Presentation® software (Version 14.1, Build 09.21.09) for presentation of the stimuli and preanalysis of the data.

Design

Out of the 22 categories, two were used to provide filler items, and ten were categories from which no members were presented in retrieval practice (i.e., Nrp categories and targets). From the other ten (Rp) categories, one member (Rp+) was practiced during retrieval practice, leaving the other member nonpracticed (Rp–). Members of the Nrp categories were divided into Nrp+ and Nrp– items, which served as baselines for the Rp+ and Rp– items, respectively. For each participant, the categories (except filler categories) were randomly assigned to category types (Nrp vs. Rp), and members of each category (except filler items) were then randomly assigned to item types (+ vs. –). Fillers were from the same categories throughout the experiment.

Procedure

The participants went through four phases of the experiment; a study, a retrieval practice, a delay, and a test phase. In the study phase, participants were shown all 44 category–member pairs once on a computer screen and were asked to remember the members with the help of the category cues. In each trial, a category word appeared to the left of the middle of the screen, together with one of the words from that category to the right of the middle of the screen. The word pair was shown for 3,000 ms, followed by a 500-ms intertrial interval (blank screen). We opted for such a short presentation of the word pairs in order to further decrease the possibility of integration of items from the same category. The study list was pseudorandomly shuffled for each participant, with the constraint that the same category could not appear within five consecutive trials. Presentation of the study list started and ended with two of the filler category–member pairs.

When the study phase was finished, participants immediately received the instructions for the retrieval practice phase. This phase consisted of three cycles. In every cycle, all Rp+ items were presented for retrieval practice once, in a random order. In each trial, the participants saw a category cue to the left of the middle of the screen and the two-letter stem of the Rp+ member of that category. The instructions were to try to recall and report the correct answer. Participants were asked to press the response button (the Enter key on the keyboard) as soon as they had the answer in mind. In order to have a valid measure of how fast an item came to mind (and not just a measure of category familiarity or feeling of knowing), we told the participants that we were curious about how fast they could recall memories, and instructed them to act as if they were on a TV quiz show, where they could lose points if they pressed the response button but could not come up with an answer immediately. After pressing the button, they were asked to type in the answer. They had 8 s to do this. If they pressed the response button or exceeded the time limit, they were shown the subsequent trial. In the first cycle, participants had 6 s to report that they knew the answer, and in the following two cycles they had 4 s. If participants did not press the response button, the next trial was introduced. The retrieval phases also started and ended with two filler trials.

After retrieval practice, the participants engaged in a 5-min two-back task, which served as a delay before the test.

The test phase consisted of 44 trials that tested memory for all of the category members. This phase also started and ended with two of the filler items. Trials were presented in the same way as in the first retrieval practice cycle, except that the category-plus-word-stem cue contained only a first-letter stem of the category member. In order to avoid output interference effects (Anderson, 2003), the test phase involved two blocks. Rp– items and their controls were tested in the first block, followed by Rp+ items and their controls in the second block. Items were randomly intermixed within both blocks. The use of different control items for the Rp+ and Rp– items was necessary in order to circumvent baseline deflation (Anderson, 2003).

Results

During analysis, we used alpha set to .05 and corrected for multiple comparisons using Bonferroni correction. The retrieval practice success rates were 85%, 86%, and 89% in the three practice cycles, respectively. The final recall performance can be seen in Table 1.

To test whether our retrieval practice manipulation was successful, we performed a one-way repeated measures analysis of variance (ANOVA) on the final recall data, with four levels of item type: Rp+, Nrp+, Rp–, and Nrp–. Item type had a significant effect on final recall, $F(3, 186) = 88.11$, $p < .001$. To test for beneficial effects of practice on the practiced items and a detrimental effect of practice on the recall of competitors, we performed two post-hoc tests. Recall of Rp+ items was significantly better than recall of their Nrp+ baseline, $t(62) = 15.31$, $p < .001$, $r = .89$, and recall of Rp– items was significantly lower than recall of their Nrp– baseline, $t(62) = 2.46$, $p = .034$ (Bonferroni corrected), $r = .30$. This shows that our item type manipulation was successful and that it provided a strong practice effect and a medium-size RIF effect.

The primary target of our investigation was the relationship between the recall RTs of Rp+ items during practice and later recall of their Rp– competitors. We analyzed first-cycle RTs only because we assumed that variance in interference, and thus in the RT data, would be greatest in the first practice cycle and would be smoothed out during further practice.

In order to rule out cheating (i.e., pressing the button when the participant did not yet know the answer), we analyzed typing time (the time that elapsed between two Enter presses: the first indicating that participants knew the response, and the second indicating that they had finished typing) for each participant. This analysis showed that no participant had individual outliers in typing times, and therefore all successfully recalled Rp+ items were included in the analysis.

Within each participant, we ranked Rp+ items by their RTs measured during the first practice cycle. Then, on the basis of this rank, we split Rp+ items into tertiles with fast, moderate, and slow RTs. For each tertile, we calculated the recall rates of the corresponding Rp– items at the final test (see Fig. 1).

To test which of the Rp– tertiles contributed to the RIF, we ran an ANOVA on the final recall data with four levels of item type (Rp–$_{1.tert}$, Rp–$_{2.tert}$, Rp–$_{3.tert}$, and Nrp–). In this analysis, Mauchly's test of sphericity was significant, Mauchly's $W = .81$; therefore, degrees of freedom were corrected using Greenhouse–Geisser estimates of sphericity. Item type had a significant effect on final recall, $F(3, 186) = 2.85$, $p = .042$. Planned contrasts (Bonferroni corrected) showed that RIF was significant only for items corresponding to second-tertile Rp– items (corresponding to Rp+ items with moderate RTs), $F(1, 62) = 9.73$, $p = .006$. Rp– items corresponding to Rp+ items with fast and slow RTs also showed lower recall than baseline, but these differences were not significant: $F(1, 62) = 0.57$, $p = .99$, for Rp– items corresponding to Rp+ items with fast RTs, and $F(1, 62) = 2.28$, $p = .41$, for Rp– items corresponding to Rp+ items with slow RTs.[1]

To assess the nature of the relationship between interference and the results of interference resolution, we conducted a repeated measures ANOVA on Rp– recall, with First-Cycle Rp+ Recall RT (fast vs. moderate vs. slow) as a within-subjects factor. We found a trend toward an effect of Rp+ RTs on recall of Rp– items, $F(2, 124) = 2.17$, $p = .118$, which was due to a tendency toward a quadratic trend in the final recall data, $F(1, 62) = 3.78$, $p = .057$,

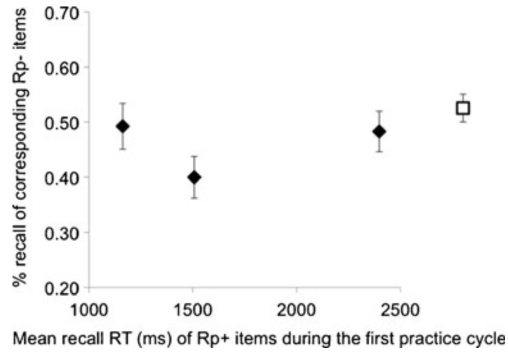

Fig. 1 Recall of Rp– items at test as a function of recall RTs for the associated Rp+ items during the first cycle of the retrieval practice phase. The empty rectangle on the right represents average Nrp– (baseline for the Rp– items) recall. The data are grouped into three tertiles according to the Rp+ RTs during the first practice cycle. Rp– recall was significantly below baseline only for second-tertile items, associated with moderate practice RTs. Error bars represent the standard errors of the means.

Table 1 Mean recall percentages at test for the four item types

	Item Type			
	Rp+	Nrp+	Rp–	Nrp–
M	.82	.45	.46	.53
SD	.17	.19	.19	.20

[1] Originally, we ran the experiment with 32 participants. In this original experiment, the final recall percentages (with standard errors in parentheses) for Rp–1.tert, Rp–2.tert, Rp–3.tert, and Nrp– items were .46 (.05), .38 (.06), .44 (.05), and .51 (.04), respectively. Rp– recall was significantly below baseline only for second-tertile items with moderate practice RTs, t (31) = 2.35, p = .038, one-tailed (Bonferroni corrected). Because this experiment was, in essence, exploratory, in order to see that this result was not a Type I error, we extended the experiment with the inclusion of another 32 participants. Logically, this was an extension rather than a replication of the original experiment (same materials, same university population, same lab, same assistant). The pattern of results obtained from this extension replicated the results of the original experiment, and the extended experiment provided greater power in detecting the same effect: Only second-tertile Rp– items were recalled below baseline, t(62) = 3.16, p = .006 (Bonferroni corrected). The data presented here are pooled from all 63 participants (as described above, one of the participant's data were excluded from the analyses).

indicating that there was one change in the direction of the relationship between Rp+ RTs during retrieval practice and the recall rate of the corresponding Rp– items at final test.

Discussion

We found practice and RIF effects with a variant of the retrieval practice paradigm that involved only two members per category. This was a novel finding, showing that the materials and design adopted here did not allow for integration of the two category members, an effect that might have masked RIF (e.g., Anderson, 2003).

Retrieval of target memories induced forgetting of competing items only when the targets were recalled with moderate retrieval RTs; RIF did not occur for competitors of memories that were recalled either too fast or too slow. This suggests that processes resolving interference during recall lead to forgetting when retrieval attempts produce moderate levels of interference.

Crucially, we showed that only a subsample of memories contribute to the RIF effect. Choosing the right sample of items to be included in the analysis might be critical for detecting the RIF effect. This can guide further investigations of the boundary conditions of RIF—for instance, when designing studies that test RIF's cue independence.

As for the exact nature of the relationship between target recall RTs and later recall of competing memories, our study is not conclusive. Our data suggest that the direction of the relationship between interference and the recall of interfering memories changes at one point from negative to positive. This would then support the suggestion that RIF is an inverted-U-shaped function of interference (Anderson & Levy, 2010; Norman et al., 2007). However, since this was supported only by a statistical tendency, strong conclusions are not warranted.

One weakness of our study is that it is hard to find three data points that would lead to rejection of a U-shaped function. A better test of this type of relationship would be to analyze final recall data binned into quartiles instead of tertiles. However, the number of items in our study was too low to provide enough power to detect such an effect if the data were split into more than three bins.[2]

Future studies could clarify several further issues raised by our results. For instance, retrieval RTs are affected not only by the magnitude of the interference that has to be resolved during retrieval, but are influenced by a range of factors, such as target strength and the strength of the associations between category cues and targets, or the relative strengths of targets and competitors. To assess the differential contribution of these factors to interference during retrieval would require new methodologies.

Another interesting issue is that the use of RT data made it impossible to analyze the effect of interference during unsuccessful retrieval attempts. Storm, Bjork, Bjork, and Nestojko (2006) showed that even unsuccessful Rp+ retrieval contributes to Rp– forgetting. Therefore, an experiment based on a measure of interference that can be collected for both retrieved and nonretrieved Rp+ items might be a significant addendum to the pattern of results presented here.

We have provided converging evidence for the interference dependence of RIF, and suggest that interference-resolving processes cause forgetting of interfering memories at moderate levels of interference. This might provide evidence for both a theoretical model based on the carryover assumption of Anderson and Levy (2010) and the computational model of Norman et al. (2007), both of which suggest that the supposed nonmonotonic function relating interference to forgetting is the sum of two linear monotonic functions: one positive, relating interference and inhibition demand, and one negative, relating interference and the success of inhibition. Although our results seem to be in line with these theories, two caveats should be mentioned here. First, as noted earlier, converging evidence will be necessary to refute either of the models describing the relationship between interference and forgetting. Second, nothing in our data suggests that the interference-resolving process involves inhibition at all. Replicating our findings with independent cues would be a strong indication of the role that inhibition plays in resolving interference.

The leap of thought introduced in the seminal article of Anderson and Bjork (1994) was the shift of attention from interference as a cause of forgetting to the consequences of interference resolution (Anderson, 2003). Our results support the view that the amount of interference plays a role in how the retrieval probabilities of related memories are shaped for later retrieval. Our study also highlights the fact that using factorial designs alone might not be sufficient to fully understand the mechanisms of interference resolution. In recent years, we have gained considerable knowledge about how the brain implements interference resolution at the systems level (e.g., Kuhl et al., 2007; Wimber, Rutschmann, Greenlee, & Bäuml, 2009). The approach and results presented here may contribute to a better understanding of interference resolution at the level of cognitive processes.

Author note This work is connected to the scientific program of the "Development of Quality-Oriented and Harmonized R+D+I Strategy and Functional Model at BME" project. This project is supported by the New Hungary Development Plan (Project ID: TÁMOP-4.2.1/B-09/1/KMR-2010-0002) by OTKA (the Hungarian National Science Foundation) Grant No. K84019. We thank Péter Pajkossy for his useful comments on a previous version of the manuscript.

[2] Binning our data into quartiles did reveal a U-shaped pattern of final recall (standard errors in parentheses): .50 (.04), .41 (.04), .44 (.04), and .48 (.04) for Rp–1.quar, Rp–2.quar, Rp–3.quar, and Rp–4.quar, respectively. Only the second- and third-quartile Rp– items differed significantly from Nrp– recall ($p = .013$ for Rp–2.quar, and $p = .042$ for Rp–3.quar). However, these comparisons do not survive Bonferroni correction. Also, this analysis did not have enough power to detect a quadratic trend in the data [$F(1, 62) = 1.98$, $p = .164$]. We thank one anonymous reviewer for drawing our attention to this problem.

Appendix

Table 2 English translations of the categories and target words used in the experiment

Category Cue	Target Member
bird	gull
bird	pelican
body part	elbow
body part	front
cloth	gloves
cloth	pajamas
drink	hot chocolate
drink	lemonade
fish	herring
fish	trout
flower	geranium
flower	lily
fruit	ananas
fruit	prune
game	dominoes
game	hide-and-seek
illness	allergy
illness	cold
insect	butterfly
insect	tick
job	model
job	soldier
kitchen utensils	microwave
kitchen utensils	whisk
mammal	bear
mammal	elephant
material	aluminium
material	concrete
musical instrument	harp
musical instrument	synthesizer
nature	cliff
nature	swamp
office utensil	calculator
office utensil	xerox
spice	marjoram
spice	parsley
sports	horse riding
sports	triathlon
weapon	rifle
weapon	spear
country[a]	Argentina[a]
country[a]	Bulgaria[a]
tools[a]	pliers[a]
tools[a]	screwdriver[a]

The original experimental materials were in Hungarian. [a] Filler categories and filler items.

References

Anderson, M. C. (2003). Rethinking interference theory: Executive control and the mechanisms of forgetting. *Journal of Memory and Language, 49,* 415–445. doi:10.1016/j.jml.2003.08.006

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

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. doi:10.1037/0278-7393.20.5.1063

Anderson, M. C., & Levy, B. J. (2010). On the relationhip between interference and inhibition in cognition. In A. S. Benjamin (Ed.), *Successful remembering and successful forgetting: A festschrift in honor of Robert A. Bjork* (pp. 107–132). New York: Psychology Press.

Anderson, M. C., & McCulloch, K. C. (1999). Integration as a general boundary condition on retrieval-induced forgetting. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 25,* 608–629. doi:10.1037/0278-7393.25.3.608

Aslan, A., Bäuml, K.-H., & Pastötter, B. (2007). No inhibitory deficit in older adults' episodic memory. *Psychological Science, 18,* 72–78. doi:10.1111/j.1467-9280.2007.01851.x

Bäuml, K.-H., Pastötter, B., & Hanslmayr, S. (2010). Binding and inhibition in episodic memory—Cognitive, emotional, and neural processes. *Neuroscience and Biobehavioral Reviews, 34,* 1047–1054. doi:10.1016/j.neubiorev.2009.04.005

Blaxton, T. A., & Neely, J. H. (1983). Inhibition from semantically related primes: Evidence of a category-specific inhibition. *Memory & Cognition, 11,* 500–510. doi:10.3758/BF03196987

BME Média és Szociológia Tanszék–Média Oktatási és Kutató Központ (BME-MOKK). (2003). *Frequency dictionary of the Hungarian Webcorpus* [Tab-separated text files]. Retrieved from http://mokk.bme.hu/resources/webcorpus

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. doi:10.1037/0278-7393.33.5.950

Halácsy, P., Kornai, A., Németh, L., Rung, A., Szakadát, I., & Trón, V. (2004). Creating open language resources for Hungarian. In M. T. Lino, M. F. Xavier, F. Ferreira, R. Costa, & R. Silva (Eds.), *Proceedings of the 4th International Conference on Language Resources and Evaluation, LREC2004* (pp. 1201–1204). Paris: European Language Resources Association.

Kornai, A., Halácsy, P., Nagy, V., Oravecz, C., Trón, V., & Varga, D (2006). Web-based frequency dictionaries for medium density languages. In *Proceedings of the 2nd International Workshop on Web as Corpus, ACL-06* (pp. 1–9). Retrieved from http://people.mokk.bme.hu/%7Ekornai/Papers/webcorp.pdf

Kuhl, B. A., Dudukovic, N. M., Kahn, I., & Wagner, A. D. (2007). Decreased demands on cognitive control reveal the neural processing benefits of forgetting. *Nature Neuroscience, 10,* 908–914.

Levy, B. J., McVeigh, N. D., Marful, A., & Anderson, M. C. (2007). Inhibiting your native language: The role of retrieval-induced forgetting during second-language acquisition. *Psychological Science, 18,* 29–34. doi:10.1111/j.1467-9280.2007.01844.x

Logan, G. D., & Cowan, W. B. (1984). On the ability to inhibit thought and action: A theory of an act of control. *Psychological Review, 91,* 295–327. doi:10.1037/0033-295X.91.3.295

Norman, K. A., Newman, E. L., & Detre, G. (2007). A neural network model of retrieval-induced forgetting. *Psychological Review, 114,* 887–953.

Perfect, T. J., Stark, L.-J., Tree, J. J., Moulin, C. J. A., Ahmed, L., & Hutter, R. (2004). Transfer appropriate forgetting: The cue-dependent nature of retrieval-induced forgetting. *Journal of Memory and Language, 51,* 399–417. doi:10.1016/j.jml.2004.06.003

Racsmány, M., & Conway, M. A. (2006). Episodic inhibition. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 32,* 44–57. doi:10.1037/0278-7393.32.1.44

Rajaram, S., Srinivas, K., & Travers, S. (2001). The effects of attention on perceptual implicit memory. *Memory & Cognition, 29,* 920–930. doi:10.3758/BF03195754

Saunders, J., & MacLeod, M. D. (2006). Can inhibition resolve retrieval competition through the control of spreading activation? *Memory & Cognition, 34,* 307–322. doi:10.3758/BF03193409

Storm, B. C., Bjork, E. L., Bjork, R. A., & Nestojko, J. F. (2006). Is retrieval success a necessary condition for retrieval-induced forgetting? *Psychonomic Bulletin & Review, 13,* 1023–1027. doi:10.3758/BF03194002

Tipper, S. P. (1985). The negative priming effect: Inhibitory priming by ignored objects. *Quarterly Journal of Experimental Psychology, 37A,* 571–590.

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,* 315–318. doi:10.1037/0278-7393.30.2.315

Verde, M. F., & Perfect, T. J. (2011). Retrieval-induced forgetting in recognition is absent under time pressure. *Psychonomic Bulletin & Review, 18,* 1166–1171. doi:10.3758/s13423-011-0143-4

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

Wimber, M., Rutschmann, R. M., Greenlee, M. W., & Bäuml, K.-H. (2009). Retrieval from episodic memory: Neural mechanisms of interference resolution. *Journal of Cognitive Neuroscience, 21,* 538–549. doi:10.1162/jocn.2009.21043