Research Article

Psychological Science
1–15

DOI: 10.1177/0956797619859531
www.psychologicalscience.org/PS

Forgetting Is a Feature, Not a Bug: Intentionally Forgetting Some Things Helps Us Remember Others by Freeing Up Working Memory Resources

TC

Vencislav Popov[1], **Ivan Marevic**[2], **Jan Rummel**[2], and **Lynne M. Reder**[1]
[1]Department of Psychology, Carnegie Mellon University and [2]Department of Psychology, Heidelberg University

Abstract
In the present study, we used an item-method directed-forgetting paradigm to test whether instructions to forget or remember one item affect memory for subsequently studied items. In two experiments (*N*s = 138 and 33, respectively), recall was higher when a word pair was preceded during study by a to-be-forgotten word pair. This effect was cumulative: Performance increased when more preceding study items were to be forgotten. The effect decreased when memory was conditioned on instructions for items appearing farther back in the study list. Experiment 2 used a dual-task paradigm that suppressed, during encoding, verbal rehearsal or attentional refreshing. Neither task removed the effect, ruling out that rehearsal or attentional borrowing is responsible for the advantage conferred from previous to-be-forgotten items. We propose that memory formation depletes a limited resource that recovers over time and that to-be-forgotten items consume fewer resources, leaving more resources available for storing subsequent items. A computational model implementing the theory provided excellent fits to the data.

Keywords
directed forgetting, item method, directed-forgetting aftereffects, computational modeling, open data, open materials, preregistered

Received 6/6/18; Revision accepted 5/14/19

Associative memory formation is an effortful process that can be disrupted by reduced study time (Malmberg & Nelson, 2003), divided attention (Craik, Govoni, Naveh-Benjamin, & Anderson, 1996), or instructions to forget (Bjork, 1972). The probability of forming associative memories decreases with stimulus difficulty; for example, recall and associative recognition are worse for low-frequency compared with high-frequency words (e.g., Criss, Aue, & Smith, 2011; Hulme, Stuart, Brown, & Morin, 2003), and the presence of low-frequency words on a study list hurts memory for other items from the same list (Diana & Reder, 2006; Malmberg & Murnane, 2002; Popov, So, & Reder, 2019; Watkins, LeCompte, & Kim, 2000). The ability to form long-term associative memories also depends on working memory (WM) capacity (Marevic, Arnold, & Rummel, 2018; Unsworth & Spillers, 2010). To explain results like these, we have proposed that binding in memory depletes a limited WM resource that recovers over time (Popov & Reder, in press; Reder, Liu, Keinath, & Popov, 2016; Reder, Paynter, Diana, Ngiam, & Dickison, 2007; Shen, Popov, Delahay, & Reder, 2018). According to this model, processing weaker items requires more resources than processing stronger items. Greater demands on limited WM resources means that there are fewer resources available to process additional items. Because the resources

Corresponding Author:
Vencislav Popov, Carnegie Mellon University, Department of Psychology, 5000 Forbes Ave., Baker Hall, 345F, Pittsburgh, PA 15213-3815
E-mail: vencislav.popov@gmail.com

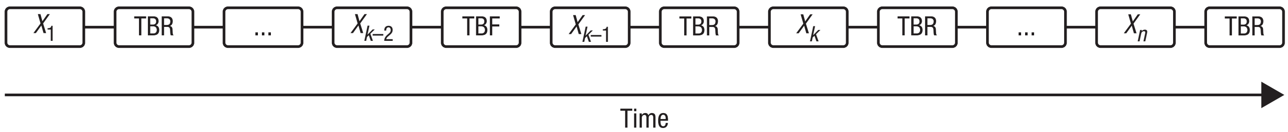

Fig. 1. Order of items during study. X denotes an item, and k denotes the position of that item in the stream (e.g., the X_{k-2} item appeared two items ago). Participants were instructed whether each item was to be remembered (TBR) or to be forgotten (TBF).

recover over time, weaker items within a list especially hurt memory for subsequent items from the same list.

Here, we tested a key prediction of the theory: Memory should be higher for items that are, during study, preceded by items consuming fewer resources. We used an item-method directed-forgetting paradigm in which each study item was directly followed by either a to-be-forgotten (TBF) or a to-be-remembered (TBR) instruction, indicating whether it would be tested later (Bjork, 1972; Golding & MacLeod, 1998). Previous studies showed worse TBF than TBR item recall (i.e., a directed-forgetting effect), but it is unknown whether memory differs for items that follow a TBR or a TBF item (i.e., a directed-forgetting aftereffect). Investigating the aftereffects of memory instructions can shed new light on the role of WM resources for long-term storage.

In line with resource-depletion-and-recovery theory (Popov & Reder, in press), our proposal is that before the instructions to remember and forget appear, participants process each item similarly, spending a proportion of their existing resources. After instructions are presented, participants continue resource-demanding processing only of TBR but not of TBF items. As a result, fewer resources remain to process items that follow one or more TBR items (compared with one or more TBF items; for an illustration of this prediction, see Fig. S3 in the Supplemental Material available online).

Early *list-method* directed-forgetting research instructing participants to forget a study list before studying a second one supports this idea by showing memory costs for the first list but memory benefits for the second list (Bjork, 1970; Epstein, 1972). List-method directed-forgetting accounts differ regarding the assumed causes for directed-forgetting costs—for example, mental context shifts (Lehman & Malmberg, 2013; Sahakyan & Kelley, 2002) and context inhibition (Pastötter, Tempel, & Bäuml, 2017). Most accounts agree, however, that directed-forgetting benefits arise because participants do not rehearse the preceding TBF list while processing the second list. Yet different mechanisms might underlie forgetting in the list-method and item-method directed-forgetting paradigms (Basden, Basden, & Gargano, 1993; Rummel, Marevic, & Kuhlmann, 2016), and it is an open question whether similar beneficial directed-forgetting aftereffects would occur on an item-by-item level. Investigating item-method directed-forgetting aftereffects allowed us to further relate the two paradigms and also to characterize this phenomenon with greater detail.

The resource-depletion-and-recovery theory makes several predictions concerning directed-forgetting aftereffects. Consider Figure 1, which depicts a study-item sequence. We predicted that memory for item X_k, $P(X_k)$, will depend on the memory instruction for the preceding items X_{k-i}, where k denotes the position of the current item, and i denotes the lag to the preceding item (e.g., the X_{k-2} item appeared two items ago). Specifically, (a) $P(X_k)$ will be higher when X_{k-1} is TBF rather than TBR; (b) these effects should be cumulative—the more preceding items are TBF, the higher $P(X_k)$ will be—and (c) these effects will also depend on the lag, i, between study items—the instruction-type effect will be greater for X_{k-1} than for X_{k-2}, and so on.

We tested these predictions in two experiments. The first experiment involved a reanalysis of the data collected by Marevic et al. (2018); the second experiment involved new data from a dual-task experiment that was designed to test whether suppressing rehearsal or dividing attention while concurrently performing the item-method directed-forgetting task would negate directed-forgetting aftereffects. To show that resource-depletion-and-recovery theory can capture the precise quantitative pattern, we also fitted a computational implementation of the account to the data.

Experiment 1: Reanalysis of the Marevic et al. (2018) Study

Method

These methods were described by Marevic et al. (2018) but are also included here to facilitate comprehension of the new information we report. The data, materials, and analysis code for the current analysis are available at osf.io/5qd94.

Participants. We recruited 138 students from Heidelberg University (110 female; age: $M = 21.96$ years, range = 19–34) who received course credit or monetary compensation for their participation. We used the full data set collected by Marevic et al. (2018), who determined the sample size so that it would allow for informative Bayesian decisions regarding the research questions tackled in their study.

Materials. A set of 96 nouns of medium frequency was drawn from the *dlex* database (Heister et al., 2011). Words were randomly paired and assigned to two sets with 24 word pairs each. One set was used in an initial practice phase and the other was used in the experimental phase. To control for item-specific effects, we counterbalanced the assignment of word-pair sets to phases. In each block, half of the word pairs were followed by instructions to forget them (TBF word pairs) and half by instructions to remember them (TBR word pairs).

Procedure. Experimental sessions started with a WM task (not analyzed here but reported by Marevic et al., 2018) and a practice phase in which participants studied 24 TBR and TBF word pairs. Participants were told to remember only the TBR word pairs for a later test and to forget the TBF word pairs. Each word pair was presented for 7 s in the center of the screen, followed by either a TBR or a TBF instruction for 2 s (i.e., the word "remember" or "forget" in German). Trials were separated by a 250-ms interstimulus interval. After all word pairs had been presented, participants solved math problems for 30 s before completing a free-recall test. In the practice phase, the free-recall test was followed by a cued-recall test for TBR items only. Recall cues were presented in random order for each participant. This practice phase was intended to familiarize participants with the paradigm and to increase their belief that the instruction to forget was genuine. However, for the real task phase, the procedure was modified so that participants were, again, presented with TBF and TBR items but were asked to recall as many TBR and TBF items as possible in the subsequent free- and cued-recall tests. Finally, participants performed another WM task (not reported) and then were debriefed and compensated for their participation.

Data analysis. We employed Bayesian statistics for the new analyses of Marevic et al.'s (2018) behavioral data. This approach had several advantages (Wagenmakers, Morey, & Lee, 2016), but most important to us was that Bayes factors (BFs) enabled us to quantify the evidence in favor of the null as well as the alternative hypotheses. We calculated BFs using bridge sampling for comparing models that included the effect of interest with models that did not. A BF close to 1 means that both models are equally likely, a BF greater than 3 is conventionally interpreted as moderate evidence, and a BF greater than 10 provides strong evidence in favor of the preferred model (Lee & Wagenmakers, 2013). We applied multilevel logistic Bayesian regressions as implemented in the *brms* package in the R programming environment (Bürkner, 2017; R Core Team, 2019), in which we included crossed random intercepts for participants and items as well as random participant slopes for directed-forgetting effects and aftereffects. The population-level regression coefficients had a weakly informative Student's t distribution prior that was zero-centered with 3 degrees of freedom and a scale of 2.5 (Gelman, Jakulin, Pittau, & Su, 2008). For the free-recall analysis, words were coded as correctly recalled when both items of a pair were recalled. All models were run with 10,000 iterations, and half of those iterations were used as burn-in. Convergence was assessed using the potential scale-reduction factor $\hat{R}$. For all parameters, $\hat{R}$ was less than 1.01, indicating good convergence.

For each item, we coded whether a TBR or TBF item preceded it. Given that the first item of a study sequence had no predecessor, it was not analyzed. To measure the cumulative effect of successive cues, we also coded how many consecutive TBR or TBF items preceded each item. We used a coding scheme that varied from −3 (3 or more consecutive TBF items preceded the current item) to +3 (3 or more consecutive TBR items preceded the current item). For example, if the current study item were preceded by a TBF and a TBR item, in that order, it would have been scored as −1 because there was only one immediately preceding TBF item. Finally, we also looked at the effect of the instructions at each lag individually without considering other potential intervening items. The output files from the *brms* analyses are available on the Open Science Framework at https://osf.io/5qd94/ in the folder labeled "analysis_output."

Results

Main effect of preceding item type. Figures 2a and 2d plot cued- and free-recall accuracy as a function of the instructions given for the current item and the preceding item. There was a directed-forgetting aftereffect; both cued recall and free recall were greater for items that were preceded by TBF items than for those preceded by TBR items (BF_{cued} = 474 and BF_{free} = 3,557 for the cued- and free-recall models with current and preceding instruction type vs. the null model with only current type). There was no interaction between instructions for the preceding item and those for the current item (BF_{cued} = 4.43 and BF_{free} = 17.77 for the cued- and free-recall models with main effects only vs. the model with an interaction).

Cumulative effect of the number of consecutive preceding TBF or TBR items. Figures 2b and 2e show cued-recall and free-recall accuracy as a function of the number of consecutive preceding TBF or TBR items. Both cued- and free-recall performance for the current item was higher when it was preceded by a greater number of consecutive TBF items and lower when it was preceded by a greater number of consecutive TBR items. The model

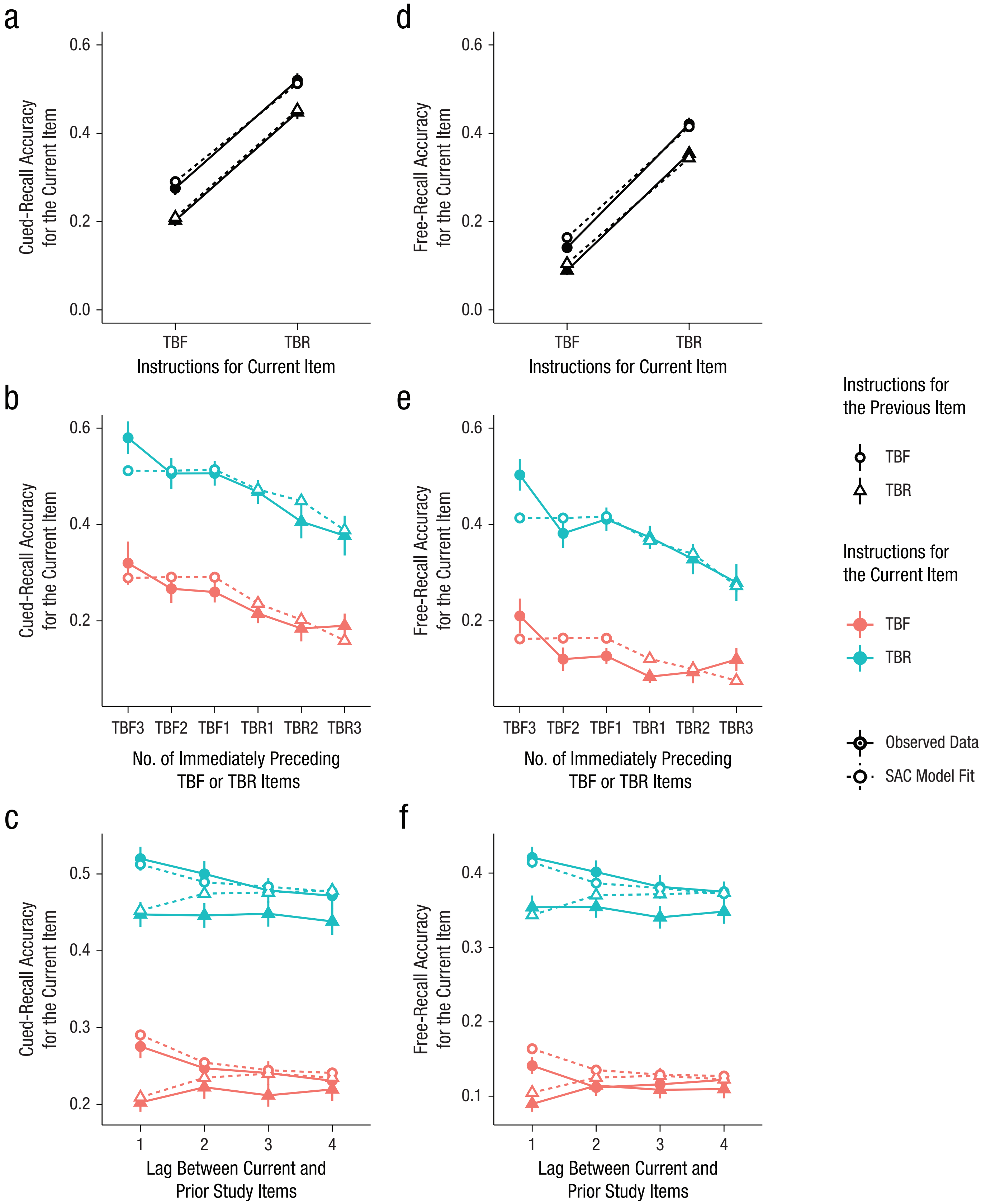

Fig. 2. Results of Experiment 1 and fit of the source-of-activation-confusion (SAC) model: cued recall (left column) and free recall (right column) for the current item, depending on (a, d) whether it was a to-be-remembered (TBR) or to-be-forgotten (TBF) item and whether it was preceded during study by a TBR or a TBF item, (b, e) how many of the immediately preceding items during study were TBR or TBF, and (c, f) the study-position lag between the current item and the prior item (e.g., how many trials ago the previous item occurred). Error bars represent ±1 *SE*. Solid points and lines represent the data; the empty points and dashed lines represent the predictions of the SAC model.

Table 1. Parameter Estimates for the Bayesian Mixed-Effects Logistic Regression on Cued Recall in Experiment 1

Effect type and predictor	Parameter estimate	Odds ratio	Bayes factor
Fixed effect			
Intercept (TBF instructions)[a]	β = −0.88	0.41 [0.30, 0.58]	
TBR instructions for the current item[a]	β = 1.17	3.21 [2.61, 3.93]	BF^ = 4.41 × 10^{32}
TBR instructions for the item at Lag 1	β = −0.41	0.66 [0.55, 0.81]	BF^ = 277
TBR instructions for the item at Lag 2	β = −0.26	0.77 [0.64, 0.93]	BF^ = 3.84
TBR instructions for the item at Lag 3	β = −0.23	0.80 [0.66, 0.96]	BF^ = 2.61
TBR instructions for the item at Lag 4	β = −0.13	0.88 [0.73, 1.05]	BF^ = 0.16
Participant random effect			
Intercept	σ = 0.79 [0.63, 0.97]		
TBR instructions for the current item[a]	σ = 0.50 [0.12, 0.78]		
TBR instructions for the item at Lag 1	σ = 0.33 [0.02, 0.68]		
Item random effect			
Intercept	σ = 0.47 [0.32, 0.68]		
Parameter comparison			
Lag 1 < Lag 2			BF+ = 7.10
Lag 2 < Lag 3			BF+ = 1.41
Lag 3 < Lag 4			BF+ = 3.63

Note: The parameter estimates reflect the means of the posterior distribution. Values in brackets are 95% Bayesian credible intervals. BF^ refers to the Bayes factor for the model that includes the parameter versus a model that does not; BF+ refers to Bayes factor evidence for the difference between the directed-forgetting aftereffect at different lags. Participants were instructed whether the current item or the items at lag *i* were to be remembered (TBR) or to be forgotten (TBF).
[a]The reference category for this analysis was TBF instruction, so the parameter estimates of the memory-instruction effects reflect the odds for correct recall with TBR instructions.

including the current item's instructions and the number of consecutive TBF or TBR preceding items fit the data better than the null model that included only the current item's instructions as a predictor (BF = 685 for cued recall and BF = 977 for free recall). There was strong evidence that the directed-forgetting effect and the directed-forgetting aftereffect did not interact (BF_{cued} = 111 and BF_{free} = 100 in favor of the cued- and free-recall models with main effects only vs. the model with an interaction term).

Interaction between preceding item type and study-position lag. Finally, Figures 2c and 2f plot cued-recall and free-recall accuracy, respectively, as a function of the preceding item type and the lag between that preceding item and the current item on the study list (i.e., ignoring the type for the intervening items). The plots clearly show that the directed-forgetting aftereffect interacted with the lag between the current item and the preceding item; the immediately preceding item had a stronger effect than the item two trials before, which in turn had a stronger effect than the item three trials before. We compared the full model, which included the instructions for items at Lags 1, 2, 3, and 4, with identical models without the factor of interest. The posterior parameter estimates from the final model and the corresponding BFs are reported in Table 1 for the cued-recall test and Table 2 for the free-recall test. The directed-forgetting aftereffect from Lag 1 was greater than the directed-forgetting aftereffect from Lag 2 for both the cued-recall and free-recall tests, and the aftereffect from Lag 3 was greater than the one from Lag 4 for the cued-recall test (for parameter comparisons, see Tables 1 and 2).

Table 2. Parameter Estimates for the Bayesian Mixed-Effects Logistic Regression on Free Recall in Experiment 1

Effect type and predictor	Parameter estimate	Odds ratio	Bayes factor
Fixed effect			
Intercept (TBF instructions)[a]	β = −1.95	0.14 [0.10, 0.20]	
TBR instructions for the current item[a]	β = 1.58	4.88 [6.82, 6.26]	BF^ = 3.52×10^{82}
TBR instructions for the item at Lag 1	β = −0.49	0.61 [0.48, 0.77]	BF^ = 397
TBR instructions for the item at Lag 2	β = −0.19	0.83 [0.67, 1.02]	BF^ = 0.63
TBR instructions for the item at Lag 3	β = −0.22	0.80 [0.65, 0.99]	BF^ = 0.78
TBR instructions for the item at Lag 4	β = −0.19	0.83 [0.67, 1.02]	BF^ = 0.20
Participant random effect			
Intercept	σ = 0.30 [0.03, 0.56]		
TBR instructions for the current item[a]	σ = 0.46 [0.10, 0.73]		
TBR instructions for the item at Lag 1	σ = 0.46 [0.06, 0.82]		
Item random effect			
Intercept	σ = 0.34 [0.19, 0.53]		
Parameter comparison			
Lag 1 < Lag 2			BF+ = 40.32
Lag 2 < Lag 3			BF+ = 0.69
Lag 3 < Lag 4			BF+ = 1.45

Note: The parameter estimates reflect the means of the posterior distribution. Values in brackets are 95% Bayesian credible intervals. BF^ refers to the Bayes factor for the model that includes the parameter versus a model that does not; BF+ refers to Bayes factor evidence for the difference between the directed-forgetting aftereffect at different lags. Participants were instructed whether the current item or the items at lag *i* were to be remembered (TBR) or to be forgotten (TBF).
[a]The reference category for this analysis was TBF instruction, so the parameter estimates of the memory-instruction effects reflect the odds for correct recall with TBR instructions.

Source-of-activation-confusion (SAC) computational model of results. Figure 2 also shows the fit of the SAC resource-depletion-and-recovery model. A full description of the model is available in the Supplemental Material and in Popov and Reder's (in press) article; we describe it only briefly and note which of the model assumptions were specifically adapted for this study.

Our model posits that semantic, episodic, and contextual information is represented as a network of interconnected nodes that vary in strength. Each node has a current activation value that increases when a node is perceived or when it receives activation from other nodes. This activation decays with time according to an exponential law to a base-level strength of the node. The base-level strength also increases with experience and decreases with time, according to a power law. When new information is studied, two processes occur. First, the current and the base-level activation values of the preexisting concept nodes are increased. Second, if this is the first occurrence of the study event, a new event node is created, and it gets associated with the corresponding concept and context nodes. If, however, the study event has occurred previously, the existing event node and its links associated with the concept and context nodes are strengthened instead.

During cued recall, the activation of the list-context node and the cue-word-concept node is raised, which then spreads activation to all nodes to which they are connected. The amount of activation that is spread from a node to any given association is multiplied by the strength of its association and divided by the sum total strength of all associated links that emanate from that node. If the current activation of an event node that is connected to the cue-concept node surpasses a retrieval threshold, then the correct target word is recalled. The model was not designed to model free recall; however, we simulated free recall by providing only the context node as a cue and evaluating the activation level of all items simultaneously. We also assumed that there would be output interference during free recall, which we simulated by exponentiating the activation values; this resulted in squashing the activation of weak items compared with stronger items.

The model also includes a resource pool that is used every time a node is retrieved, created, or strengthened. The resource cost of strengthening a node is equal to the degree to which a node is strengthened. Similarly, the resource cost of retrieving a node is equal to the amount of activation necessary to reach the retrieval threshold. During study, if the currently available resource pool is sufficient for storing an item, the memory trace is built or strengthened by the default learning rate. However, if there are currently fewer resources available than required, the memory trace is strengthened proportionally to the remaining resources. The resource pool recovers at a linear rate until it reaches the maximum WM resource capacity.

For the current experiment, we assumed that when an item appears, an episode node is created with a default base-level strength, regardless of the instruction type. Then, when the instruction appears, the episode node for TBR items is strengthened again, whereas the node for TBF items is not. We fitted the model by simulating data for each participant, given the specific trial sequence. Six parameters were optimized by minimizing the root-mean-square error (RMSE) of the cued-recall and free-recall data averaged over all participants, the current instruction type, and the number of consecutive preceding TBR or TBF items (24 data points; see Figs. 2b and 2e). In our initial modeling, we estimated separate learning rates for the strengthening during item and instruction presentation. These two estimates were roughly equal and the model did not fit the data significantly better than the simpler model with a single learning rate for the strengthening during both item and instruction presentation. The final model parameters consisted of a learning rate (δ) of 0.553, which governs how much the base-level strength of nodes is increased with each exposure; a resource-recovery rate (w_r) of 0.526; retrieval thresholds (θ) of 0.219 for cued recall and 0.167 for free recall; and standard deviation of the activation noise (σ) of 0.831 for cued recall and 0.431 for free recall. All remaining parameters had the default values we have used in prior models. The model provided very good fits to the cued-recall data (RMSE = 0.026, R^2 = .963) and free-recall data (RMSE = 0.034, R^2 = .944). It is noteworthy that the model also captured the interaction between instruction type and lag (see Figs. 2c and 2f), although the parameters were not optimized to fit those data points.

Experiment 2

Despite good model fit, there were still alternative explanations for Experiment 1's results. People may rehearse or reactivate the memory traces of preceding items while processing the current item (Camos, Lagner, & Barrouillet, 2009; McFarlane & Humphreys, 2012). Such rehearsal or attentional borrowing is more likely when the preceding item was TBR rather than TBF (Bjork, 1970), resulting in diminished processing for the current item. Similarly, the retrieving-effectively-from-memory (REM) model (Shiffrin & Steyvers, 1997; Lehmann & Malmberg, 2013) postulates that there is a limited rehearsal buffer and that memory-trace strength depends on how much of the buffer is currently available. The REM model would attribute the directed-forgetting aftereffect to the fact that TBF items are not rehearsed, which frees buffer space for the rehearsal of the current item.

In Experiment 2, we tested whether suppressing rehearsal during study would eliminate the directed-forgetting aftereffect to rule out that it is due to greater rehearsal of preceding TBR items (for a similar argument concerning the effect of articulatory suppression on rehearsal-based explanations for the regular directed-forgetting effect, see Hourihan, Ozubko, & MacLeod, 2009). We further tested whether the directed-forgetting aftereffect would be attenuated under divided attention to rule out that it is due to allocating attention to previous pairs (attentional refreshing) instead of the current pair (for illustrations, see Figs. S4 and S5 in the Supplemental Material). A stable directed-forgetting aftereffect under suppressed rehearsal or divided attention would support the resource-depletion-and-recovery explanation.

Method

The rationale, method, and original analysis plan for this experiment were preregistered on the Open Science Framework (https://osf.io/yugkt). We deviated from the preregistered analysis plan after realizing that a Bayesian analysis of variance (ANOVA) would not be appropriate for analyzing proportion data; instead, we conducted a Bayesian logistic regression. The parametric predictions were not included in the preregistration plan. This makes them exploratory for Experiment 1 but confirmatory for Experiment 2. The data, materials, and analysis code for this experiment are available at osf.io/5qd94.

Table 3. Counterbalancing Order for the Four Conditions in Experiment 2 According to a Balanced Latin-Square Design

Order	Blocks 1 and 2	Blocks 3 and 4	Blocks 5 and 6	Blocks 7 and 8
1	Rehearsal suppression	Divided attention	Rehearsal suppression + divided attention	Control
2	Divided attention	Control	Rehearsal suppression	Rehearsal suppression + divided attention
3	Control	Rehearsal suppression + divided attention	Divided attention	Rehearsal suppression
4	Rehearsal suppression + divided attention	Rehearsal suppression	Control	Divided attention

Note: Each row represents a unique order, ensuring that each secondary task was followed and preceded by each other condition at least once. Secondary tasks of the same type were always grouped in two consecutive blocks. The control condition had no secondary task.

Participants. Course credit or monetary compensation were given to 33 students from Heidelberg University (22 female; age: M = 22.36 years, range = 18–31) who participated in individual sessions. We preregistered this experiment with sample-size requirements of at least 16 participants based on a priori considerations of statistical power. To have enough observations for computational-modeling approaches, we nevertheless decided to collect more data before we ever looked at the data. Because our initial power considerations were based on the assumption that we would conduct a 2 × 4 ANOVA, they were also not compatible with the Bayesian logistic regression that we used for the final analysis. However, all BFs that we calculated provided clear evidence in favor of either the alternative or the null hypothesis, implying that the present sample size was large enough to allow for meaningful conclusions from the data.

Materials. We selected 448 words of medium frequency from the *dlex* database (Heister et al., 2011) and randomly paired them to form 224 word pairs. The task was divided into eight blocks. Each block consisted of 12 TBF and 12 TBR word pairs. The memory instructions for individual item pairs were presented in random order for each participant. The first four items (two TBF, two TBR) of each block served as primacy buffers and were not included in the analyses.

Procedure. Participants first received general instructions for the directed-forgetting task asking them to remember only items that were followed by TBR instructions and to forget items followed by TBF instructions. Participants were informed that they were about to complete eight study-test blocks of this task while performing a different secondary task in each block. At the beginning of each block, the respective secondary task was explained (see below). Then, each block featured a study phase, in which 12 TBF and 12 TBR items were presented sequentially with a random permutation of the item-type order. All other aspects of the main study procedure were identical to those in Experiment 1. During study, participants performed different secondary tasks, which changed every two blocks. The order of secondary tasks was systematically varied across participants using a Latin-square design (see Table 3).

In the control blocks, no secondary task was added to the study phase. For the rehearsal-suppression blocks, participants wore headphones over which were played 60-beats-per-minute metronome sounds and were asked to say the German word “der” (the equivalent to “the” in English) aloud every time they heard the metronome. Additionally, they had to press the “J” key or “F” key whenever saying “der,” to keep the motor component equal across blocks. The assignment of keys was counterbalanced across participants. For the divided-attention blocks, even and odd two-digit numbers were continuously presented over participants’ headphones. They had to press the “J” key for even numbers and the “F” key for odd numbers (key assignment counterbalanced). A new number was presented every 2,000 ms, on average, but interstimulus intervals varied between 1,250 ms and 2,750 ms to prevent habituation. For the combined rehearsal-suppression and divided-attention task, participants were also presented with even and odd two-digit numbers but made verbal odd/even judgments. Additionally, they had to press the “J” or “F” key (counterbalanced) with each judgment to align motor demands to the other secondary tasks. The experimenter was present during the entire session and monitored compliance with the secondary task; if participants stopped performing the secondary task, the experimenter reminded them to continue.

This divided-attention task was designed to reduce the attention paid to the main task without requiring participants to remember the numbers. In contrast to the resource-depletion-and-recovery explanation, which proposes that different amounts of resources are depleted at Time t – 1, the attention-borrowing explanation implies that the effect is retroactive; that is, during the current trial at Time t, participants redirected attention back to the item presented at Time t – 1. The divided-attention task would remove the directed-forgetting aftereffect in the latter but not in the former case (for more information, see the Supplemental Material).

Following each block's study phase, participants always solved math problems for 30 s before they performed a free-recall test. For these tests, they were always asked to recall as many TBR items as possible in 2 min. We did not ask participants to recall TBF items because there were multiple study-test blocks, and thus, a TBF-item recall instruction would not have come as a surprise after the first block. Participants were specifically encouraged to recall both words of the pairs, if possible, but if they could recall only one word of the pair, they should report it as well. Then participants performed a cued-recall test in which they were presented with the first words of all TBR item pairs they had studied (in random order) and were asked to recall the second word. After four blocks, participants were given a 3-min break, in which they received water but had to stay in the laboratory. After completing all eight blocks, participants were asked whether they used a certain forgetting strategy and some demographic questions.

Results

Main effect of preceding item type and dual-task condition. Figures 3a and 3d plot cued-recall and free-recall accuracy as a function of the memory instructions for the preceding item and the dual-task condition. Both cued and free recall were higher for items that were preceded by TBF items rather than TBR items (BF_{cued} = 13 and BF_{free} = 134 for the cued- and free-recall models with dual-task condition and preceding instruction type vs. the null model with only dual-task condition as a factor). Overall, memory performance was lower in all dual-task conditions compared with the control condition (BF_{cued} = 411 and BF_{free} = 500 for the cued- and free-recall models with dual-task condition as a main factor vs. the null model). This overall memory decline indicates that the dual-task condition was effective in preventing participants from engaging in articulatory rehearsal or attentional refreshing during study. Nevertheless, the directed-forgetting aftereffect was present in all conditions because the preceding items' instructions did not interact with dual-task condition (BF_{cued} = 395 and BF_{free} = 1,515 for the cued- and free-recall models with main effects only against the models with an interaction). Because the main effect of preceding instruction type did not differ between conditions, we report all remaining analyses collapsed over conditions.

Cumulative effect of the number of consecutive preceding TBF or TBR items. Figures 3b and 3e show cued- and free-recall accuracy as a function of the number of consecutive preceding TBF or TBR items. Both cued- and free-recall performance for the current item were higher when it was preceded by a greater number of consecutive TBF items and lower when it was preceded by a greater number of consecutive TBR items. The model including the number of consecutive TBF or TBR items fitted the data better than the null model (BF_{cued} = 1,402 and BF_{free} = 99).

Interaction between preceding cue and study-position lag. Finally, the directed-forgetting aftereffect interacted with the study lag between the current item and the preceding item; the immediately preceding item had a stronger effect than the item two trials before, which in turn had a stronger effect than the item that occurred three trials before (see Figs. 3c and 3f). We compared the full model, which included the instructions for items at Lags 1, 2, 3, and 4, with identical models without the factor of interest. The posterior parameter estimates from the final model and the corresponding BFs are reported in Table 4 for the cued-recall test and Table 5 for the free-recall test.

SAC computational modeling. As in Experiment 1, we fitted the SAC model by simulating data for each participant, given the specific trial sequence. There is no rehearsal mechanism in the model, and for that reason, we ignored the dual-task conditions and modeled only the effect of the prior cue. The same six parameters were optimized by minimizing the RMSE of the cued-recall and free-recall data averaged over the number of consecutive preceding TBR or TBF items (12 data points; see Figs. 3b and 3e). In addition, we had to increase the free-recall output-interference exponent parameter to account for the different performance in free and cued recall. The estimated parameters were very similar to those of Experiment 1: The learning rate (δ) was 0.639, the resource-recovery rate (w_r) was 0.551, the retrieval thresholds (θ) were 0.279 for cued recall and 0.457 for free recall, and the standard deviations of the activation noise (σ) were 0.451 for cued recall and 0.868 for free recall. All remaining parameters had the default values that we used in prior models. The model provided excellent fits to the cued-recall data (RMSE = 0.008, R^2 = .991) and free-recall data (RMSE = 0.005, R^2 = .984). It is noteworthy that the model also captured the fact that the directed-forgetting aftereffect decreases with lag (see Figs. 3c and 3f), even

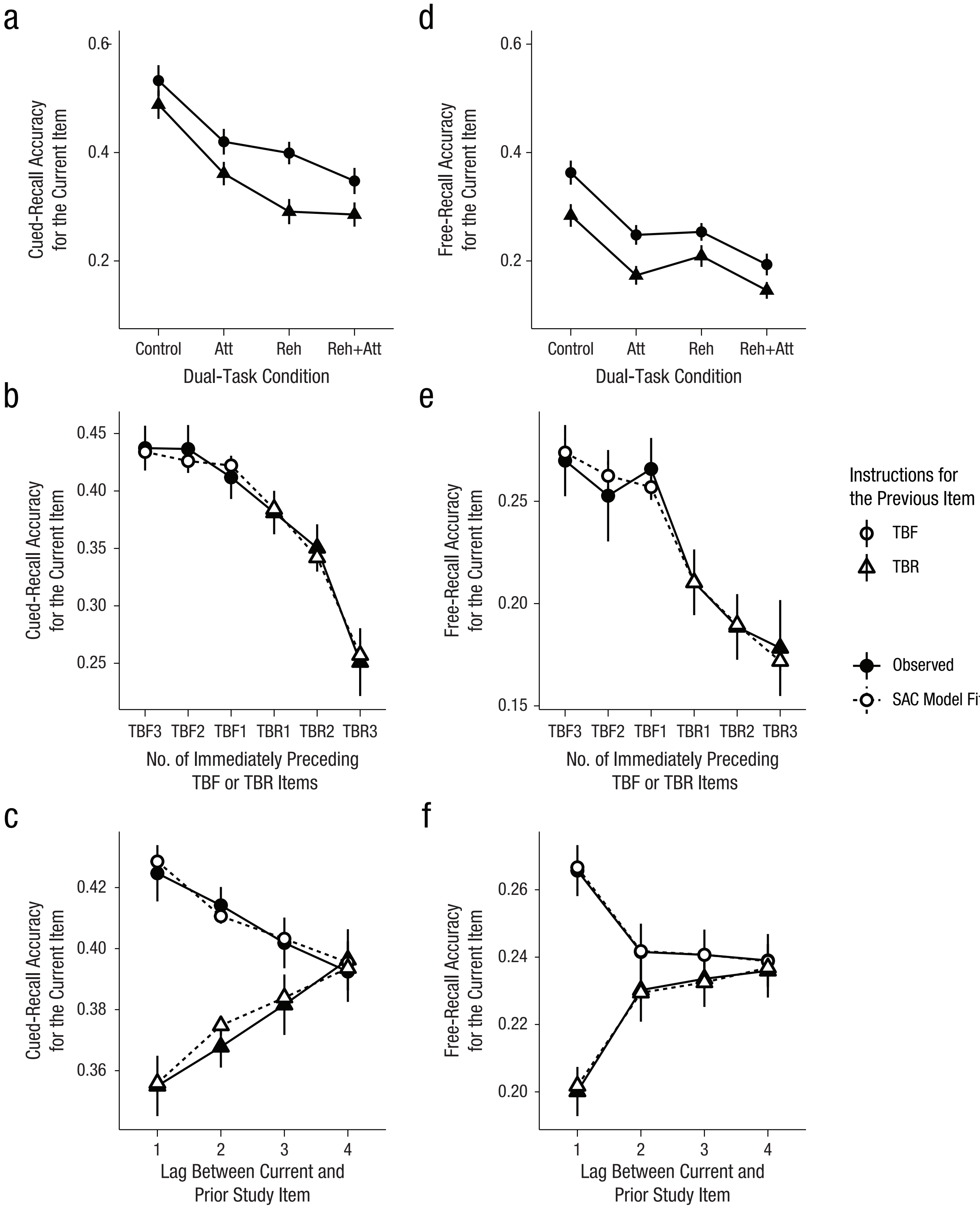

Fig. 3. Results of Experiment 2 and source-of-activation-confusion (SAC) model fits: cued recall (left column) and free recall (right column) for the current item, depending on (a, d) whether it was preceded during study by a to-be-remembered (TBR) or to-be-forgotten (TBF) item and the dual-task condition (control = no dual task, Att = divided attention, Reh = rehearsal suppression, Reh + Att = simultaneous divided attention and rehearsal suppression), (b, e) how many of the immediately preceding items during study were TBR or TBF, and (c, f) the study-position lag between the current item and the prior item (e.g., how many trials ago the previous item occurred). Error bars represent ±1 *SE*. Solid points and lines represent the data; the empty points and dashed lines represent the predictions of the SAC model.

Table 4. Parameter Estimates for the Bayesian Mixed-Effects Logistic Regression on Cued Recall in Experiment 2

Effect type and predictor	Parameter estimate	Odds ratio	Bayes factor
Fixed effect			
Intercept (TBF instructions; control)[a]	β = 0.44	1.56 [0.92, 2.67]	
Effects of dual-task condition			
Divided-attention condition	β = −0.66	0.52 [0.31, 0.87]	BF^ = 177.57
Rehearsal-suppression condition	β = −0.54	0.43 [0.26, 0.71]	BF^ = 1,874
Divided-attention condition + rehearsal-suppression condition	β = −1.13	0.32 [0.19, 0.54]	BF^ > 15 × 10^3
Effects of instructions			
TBR instructions for the item at Lag 1	β = −0.39	0.68 [0.54, 0.85]	BF^ = 17.94
TBR instructions for the item at Lag 2	β = −0.28	0.76 [0.62, 0.92]	BF^ = 2.93
TBR instructions for the item at Lag 3	β = −0.15	0.86 [0.71, 1.05]	BF^ = 0.18
TBR instructions for the item at Lag 4	β = −0.01	0.99 [0.81, 1.20]	BF^ = 0.05
Participant random effect			
Intercept (control)	σ = 1.14 [0.85, 1.52]		
Divided-attention condition	σ = 0.65 [0.19, 1.12]		
Rehearsal-suppression condition	σ = 0.56 [0.11, 1.00]		
Divided-attention condition + rehearsal-suppression condition	σ = 0.69 [0.28, 1.13]		
TBR instructions for the item at Lag 1	σ = 0.28 [0.02, 0.69]		
Item random effect			
Intercept	σ = 0.91 [0.76, 1.08]		
Parameter comparison			
Lag 1 < Lag 2			BF+ = 3.37
Lag 2 < Lag 3			BF+ = 4.65
Lag 3 < Lag 4			BF+ = 5.57

Note: The parameter estimates reflect the means of the posterior distribution. Values in brackets are 95% Bayesian credible intervals. BF^ refers to the Bayes factor for the model that includes the parameter versus a model that does not; BF+ refers to Bayes factor evidence for the difference between the directed-forgetting aftereffect at different lags. Participants were instructed whether the current item or the items at lag *i* were to be remembered (TBR) or to be forgotten (TBF).
[a]The reference category for this analysis was TBF instruction in the control condition, so the parameter estimates of the memory-instruction effects reflect the odds for correct recall with TBR instructions.

though the parameters were not optimized to fit those data points.

General Discussion

We demonstrated a novel directed-forgetting aftereffect: When an item is TBF rather than TBR, memory for the subsequent item benefits. This effect occurs in both cued and free recall and is cumulative: The more that preceding items are TBF, the higher the memory benefits; the effect decreases when memory is conditioned on instructions for items appearing further back in the study list. The directed-forgetting aftereffect was replicable and remarkably consistent across the two experiments; the cued-recall odds ratios associated with items preceded by TBR items relative to TBF items were 0.66 and 0.67, respectively.

Previous research has also shown improved memory for whole lists when a preceding list was TBF rather than TBR (Bjork, 1970; Epstein, 1972). This is, however,

Table 5. Parameter Estimates for the Bayesian Mixed-Effects Logistic Regression on Free Recall in Experiment 2

Effect type and predictor	Parameter estimate	Odds ratio	Bayes factor
Fixed effect			
Intercept (TBF instructions; control)[a]	β = −0.65	0.52 [0.34, 0.78]	
Effect of dual-task condition			
Divided-attention condition	β = −0.77	0.46 [0.31, 0.69]	BF^ > 15 × 10^3
Rehearsal-suppression condition	β = −0.65	0.52 [0.34, 0.79]	BF^ = 651.17
Divided-attention condition + rehearsal-suppression condition	β = −1.15	0.32 [0.19, 0.51]	BF^ > 15 × 10^3
Effect of instructions			
TBR instructions for the item at Lag 1	β = −0.48	0.62 [0.47, 0.81]	BF^ = 30.53
TBR instructions for the item at Lag 2	β = −0.12	0.89 [0.72, 1.10]	BF^ = 0.15
TBR instructions for the item at Lag 3	β = −0.09	0.92 [0.75, 1.13]	BF^ = 0.05
TBR instructions for the item at Lag 4	β = −0.08	0.92 [0.74, 1.14]	BF^ = 0.06
Participant random effect			
Intercept (control)	σ = 0.63 [0.42, 0.90]		
Divided-attention condition	σ = 0.21 [0.01, 0.58]		
Rehearsal-suppression condition	σ = 0.44 [0.04, 0.86]		
Divided-attention condition + rehearsal-suppression condition	σ = 0.67 [0.19, 1.18]		
TBR instructions for the item at Lag 1	σ = 0.38 [0.03, 0.77]		
Item random effect			
Intercept	σ = 0.70 [0.54, 0.87]		
Parameter comparison			
Lag 1 < Lag 2			BF+ = 69.42
Lag 2 < Lag 3			BF+ = 1.37
Lag 3 < Lag 4			BF+ = 1.04

Note: The parameter estimates reflect the means of the posterior distribution. Values in brackets are 95% Bayesian credible intervals. BF^ refers to the Bayes factor for the model that includes the parameter versus a model that does not; BF+ refers to Bayes factor evidence for the difference between the directed-forgetting aftereffect at different lags. Participants were instructed whether the current item or the items at lag *i* were to be remembered (TBR) or to be forgotten (TBF).
[a]The reference category for this analysis was TBF instruction in the control condition, so the parameter estimates of the memory-instruction effects reflect the odds for correct recall with TBR instructions.

the first study to demonstrate directed-forgetting aftereffects on an item level and to characterize in detail how the precise order of TBR and TBF items affects memory for subsequent items. The present findings indicate similarities between the two directed-forgetting methods but also provide new theoretical insight, because the item method allows for a more fine-grained investigation of directed-forgetting aftereffects. For example, researchers have argued that the list-method directed-forgetting aftereffect is due to less rehearsal borrowing (Bjork, 1970; Sahakyan & Kelley, 2002). This explanation is unlikely to hold for the item method because the directed-forgetting aftereffects in our experiments were not attenuated when rehearsal was prevented.

What causes item-method directed-forgetting aftereffects? We propose that memory formation and storage deplete a limited resource that recovers over time

(Popov & Reder, in press; Reder et al., 2007). Within this framework, TBR items deplete more resources, and they leave fewer resources for processing subsequent items. A computational model implementing the theory provided excellent fits to the cued- and free-recall data. Although we do not know whether directed-forgetting aftereffects would appear in other tasks (e.g., recognition) or with other materials (e.g., single words), directed forgetting is not the only manipulation that leads to aftereffects; similar patterns occur when the preceding items are of high rather than low frequency or have been repeated more often in the experiment (Popov & Reder, in press). These other aftereffects occur under a variety of encoding and retrieval conditions, and the general pattern is remarkably similar to the one found for directed forgetting here. Item-specific aftereffects seem to be a general mnemonic phenomenon that can be tied together with the current model.

The idea that the required processing resources differ for TBR and TBF items is not new. Fawcett and Taylor (2008, 2012) argued that participants actively withdraw attentional resources from TBF items when being presented with an instruction to forget, freeing resources to process prior TBR items. The key difference between this research and ours is that, whereas Fawcett and Taylor measured incidental memory for secondary probes not relevant to the primary memory task presented shortly after the forget instructions, we measured intentional memory for subsequent study items. Fawcett and Taylor found response times to post-TBF probes to be slower than to post-TBR probes and recognition memory for post-TBF probes to be worse than for post-TBR probes. Fawcett and Taylor (2012) suggested that these effects are indicators of greater processing in the immediate aftermath of TBF compared with TBR instructions. Our experiments were not designed to measure forget-instruction-induced attention withdrawal, and thus, our findings do not speak for or against the existence of such a process. However, if such an attention-withdrawal process existed, it would need to be of short duration and not overly resource taxing. Otherwise, we would not have observed memory benefits from preceding TBF items but, rather, we would have observed the opposite.

Are there alternative explanations for the directed-forgetting-aftereffect phenomenon? We discount three possibilities. First, the directed-forgetting aftereffect cannot be due to continued rehearsal of preceding TBR items; articulatory suppression makes verbal rehearsal nearly impossible, and it would have eliminated the effect were it due to rehearsal borrowing. Second, if memory for the current item was worse because participants were directing their attention to the preceding TBR items, then dividing attention should have reduced the directed-forgetting aftereffect proportionally to the overall reduction in memory. This prediction follows if we assume that dividing attention makes it less likely that participants use their remaining attentional resources to process preceding items but that they would rather focus them mostly on the current item (see Fig. S5 in the Supplemental Material). Whereas dividing attention reduced recall, the directed-forgetting aftereffect was not attenuated. It is nevertheless possible to imagine alternative formulations of attentional refreshing that might be consistent with these data. A final alternative is that when an item is forgotten, the surrounding items become more distinct and easier to retrieve (e.g., Brown, Neath, & Chater, 2007; Sederberg, Howard, & Kahana, 2008). This explanation would predict that TBR items should impair memory for both preceding and following study items. We did not find support for this prediction; accuracy for the current item did not differ depending on whether it was followed by TBF or TBR items during study (for details, see the Supplemental Material).

The disparity between effects of preceding and subsequent item types distinguishes the directed-forgetting aftereffect from general distinctiveness effects, in which distinct items impair memory for all surrounding items (Detterman, 1975). The fact that memory for the current item was not affected by whether the subsequent item was TBR or TBF also renders a compartmentalization explanation—for example, as suggested by the REM buffer model of Lehmann and Malmberg (2013)—less likely. Their model proposes that the presentation of distinct items causes previously studied items to be dropped from rehearsal and that distinct items are more persistent (Kamp, Lehman, Malmberg, & Donchin, 2016). A direct computational comparison of the predictions of the REM and SAC models would be necessary to adjudicate between the alternative interpretations and presents an interesting avenue for future research.

Action Editor

Charles Hulme served as action editor for this article.

Author Contributions

L. M. Reder developed the resource-depletion-and-recovery theory. V. Popov developed the predictions and the study concept. V. Popov, I. Marevic, and J. Rummel designed Experiment 2. Testing and data collection for Experiment 2 were performed by I. Marevic, and I. Marevic and V. Popov analyzed the data. V. Popov implemented the source-of-activation-confusion (SAC) model. V. Popov and I. Marevic drafted the initial manuscript. J. Rummel and L. M. Reder provided critical feedback and revisions. All the authors approved the final manuscript for submission.

ORCID iDs

Vencislav Popov https://orcid.org/0000-0002-8073-4199
Jan Rummel https://orcid.org/0000-0002-0443-1743

Acknowledgments

We are grateful to Matt So for proofreading the final manuscript.

Declaration of Conflicting Interests

The author(s) declared that there were no conflicts of interest with respect to the authorship or the publication of this article.

Supplemental Material

Additional supporting information can be found at http://journals.sagepub.com/doi/suppl/10.1177/0956797619859531

Open Practices

TC

All data and materials for both experiments have been made publicly available via the Open Science Framework and can be accessed at osf.io/5qd94. The design and analysis plans for Experiment 2 were preregistered at osf.io/yugkt. The complete Open Practices Disclosure for this article can be found at http://journals.sagepub.com/doi/suppl/10.1177/0956797619859531. This article has received the badges for Open Data, Open Materials, and Preregistration. More information about the Open Practices badges can be found at http://www.psychologicalscience.org/publications/badges.

References

Basden, B. H., Basden, D. R., & Gargano, G. J. (1993). Directed forgetting in implicit and explicit memory tests: A comparison of methods. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*, 603–616.

Bjork, R. A. (1970). Positive forgetting: The noninterference of items intentionally forgotten. *Journal of Verbal Learning and Verbal Behavior*, *9*, 255–268.

Bjork, R. A. (1972). Theoretical implications of directed forgetting. In A. W. Melton & E. Martin (Eds.), *Coding processes in human memory* (pp. 217–235). Washington, DC: Winston.

Brown, G. D., Neath, I., & Chater, N. (2007). A temporal ratio model of memory. *Psychological Review*, *114*, 539–576.

Bürkner, P.-C. (2017). brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software*, *80*(1). doi:10.18637/jss.v080.i01

Camos, V., Lagner, P., & Barrouillet, P. (2009). Two maintenance mechanisms of verbal information in working memory. *Journal of Memory and Language*, *61*, 457–469.

Craik, F. I., Govoni, R., Naveh-Benjamin, M., & Anderson, N. D. (1996). The effects of divided attention on encoding and retrieval processes in human memory. *Journal of Experimental Psychology: General*, *125*, 159–180.

Criss, A. H., Aue, W. R., & Smith, L. (2011). The effects of word frequency and context variability in cued recall. *Journal of Memory and Language*, *64*, 119–132.

Detterman, D. K. (1975). The von Restorff effect and induced amnesia: Production by manipulation of sound intensity. *Journal of Experimental Psychology: Human Learning and Memory*, *1*, 614–628.

Diana, R. A., & Reder, L. M. (2006). The low-frequency encoding disadvantage: Word frequency affects processing demands. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*, 805–815.

Epstein, W. (1972). Mechanisms of directed forgetting. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 6, pp. 147–191). San Diego, CA: Academic Press.

Fawcett, J. M., & Taylor, T. L. (2008). Forgetting is effortful: Evidence from reaction time probes in an item-method directed forgetting task. *Memory & Cognition*, *36*, 1168–1181.

Fawcett, J. M., & Taylor, T. L. (2012). The control of working memory resources in intentional forgetting: Evidence from incidental probe word recognition. *Acta Psychologica*, *139*, 84–90.

Gelman, A., Jakulin, A., Pittau, M. G., & Su, Y. S. (2008). A weakly informative default prior distribution for logistic and other regression models. *The Annals of Applied Statistics*, *2*, 1360–1383.

Golding, J. M., & MacLeod, C. M. (1998). *Intentional forgetting: Interdisciplinary approaches*. Mahwah, NJ: Erlbaum.

Heister, J., Würzner, K.-M., Bubenzer, J., Pohl, E., Hanneforth, T., Geyken, A., & Kliegl, R. (2011). dlexDB-eine lexikalische datenbank für die psychologische und linguistische forschung [dlexDB: A lexical database for psychological and linguistic research]. *Psychologische Rundschau*, *62*(1), 10–20.

Hourihan, K. L., Ozubko, J. D., & MacLeod, C. M. (2009). Directed forgetting of visual symbols: Evidence for nonverbal selective rehearsal. *Memory & Cognition*, *37*, 1059–1068.

Hulme, C., Stuart, G., Brown, G. D., & Morin, C. (2003). High- and low-frequency words are recalled equally well in alternating lists: Evidence for associative effects in serial recall. *Journal of Memory and Language*, *49*, 500–518.

Kamp, S. M., Lehman, M., Malmberg, K. J., & Donchin, E. (2016). A buffer model account of behavioral and ERP patterns in the Von Restorff paradigm. *AIMS Neuroscience*, *3*, 181–202.

Lee, M. D., & Wagenmakers, E.-J. (2013). *Bayesian cognitive modeling: A practical course*. Cambridge, England: Cambridge University Press.

Lehman, M., & Malmberg, K. J. (2013). A buffer model of memory encoding and temporal correlations in retrieval. *Psychological Review*, *120*, 155–189.

Malmberg, K. J., & Murnane, K. (2002). List composition and the word-frequency effect for recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*, 616–630.

Malmberg, K. J., & Nelson, T. O. (2003). The word frequency effect for recognition memory and the elevated-attention hypothesis. *Memory & Cognition*, *31*, 35–43.

Marevic, I., Arnold, N. R., & Rummel, J. (2018). Item-method directed forgetting and working memory capacity: A hierarchical multinomial modeling approach. *The Quarterly Journal of Experimental Psychology*, *71*, 1070–1080.

McFarlane, K. A., & Humphreys, M. S. (2012). Maintenance rehearsal: The key to the role attention plays in storage and forgetting. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *38*, 1001–1018.

Pastötter, B., Tempel, T., & Bäuml, K.-H. T. (2017). Long-term memory updating: The reset-of-encoding hypothesis in list-method directed forgetting. *Frontiers in Psychology*, *8*, Article 2076. doi:10.3389/fpsyg.2017.02076

Popov, V., & Reder, L. (in press). *Frequency effects on memory: A resource-limited theory*. *Psychological Review*. doi:10.1037/rev0000161

Popov, V., So, M., & Reder, L. (2019). *Word frequency affects binding probability not memory precision*. Retrieved from PsyArXiv Preprints: https://psyarxiv.com/deyjm/

R Core Team. (2019). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

Reder, L. M., Liu, X. L., Keinath, A., & Popov, V. (2016). Building knowledge requires bricks, not sand: The critical role of familiar constituents in learning. *Psychonomic Bulletin & Review*, *23*, 271–277.

Reder, L. M., Paynter, C., Diana, R. A., Ngiam, J., & Dickison, D. (2007). Experience is a double-edged sword: A computational model of the encoding/retrieval trade-off with familiarity. In A. S. Benjamin & B. H. Ross (Eds.), *Skill and strategy in memory use: The psychology of learning and motivation* (Vol. 48, pp. 271–312). San Diego, CA: Academic Press.

Rummel, J., Marevic, I., & Kuhlmann, B. G. (2016). Investigating storage and retrieval processes of directed forgetting: A model-based approach. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *42*, 1526–1543.

Sahakyan, L., & Kelley, C. M. (2002). A contextual change account of the directed forgetting effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*, 1064–1072.

Sederberg, P. B., Howard, M. W., & Kahana, M. J. (2008). A context-based theory of recency and contiguity in free recall. *Psychological Review*, *115*, 893–912.

Shen, Z., Popov, V., Delahay, A. B., & Reder, L. M. (2018). Item strength affects working memory capacity. *Memory & Cognition*, *46*, 204–215.

Shiffrin, R. M., & Steyvers, M. (1997). A model for recognition memory: REM—retrieving effectively from memory. *Psychonomic Bulletin & Review*, *4*, 145–166.

Unsworth, N., & Spillers, G. J. (2010). Variation in working memory capacity and episodic recall: The contributions of strategic encoding and contextual retrieval. *Psychonomic Bulletin & Review*, *17*, 200–205.

Wagenmakers, E.-J., Morey, R. D., & Lee, M. D. (2016). Bayesian benefits for the pragmatic researcher. *Current Directions in Psychological Science*, *25*, 169–176.

Watkins, M. J., LeCompte, D. C., & Kim, K. (2000). Role of study strategy in recall of mixed lists of common and rare words. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*, 239–245.