Delayed Testing in Directed Forgetting Dissociates Active and Passive Forms of Forgetting

Jonathon Whitlock[1], Huiyu Ding[1], Ryan Hubbard[1,2], Lili Sahakyan[1,2]

1. Department of Psychology, University of Illinois at Urbana-Champaign
2. Beckman Institute for Advanced Science and Technology

Correspondence concerning this article should be addressed to Jonathon Whitlock, Department of Psychology, University of Illinois at Urbana-Champaign, 603 E. Daniel St., Champaign, Illinois 61820. E-mail: jsw6@illinois.edu

Abstract

Across two experiments, we assessed the rates of relative forgetting following instructions to remember or forget information in an item-method directed forgetting paradigm across several retention intervals. In addition to the Forget and Remember cues, we also included Thought Substitution (TS) cues in the same design instructing participants to mentally shift to a different context on some study trials. TS cues have been shown to impair memory compared to Remember cues, but not as effectively as cues to Forget in item-method studies (Hubbard & Sahakyan, 2021). The results demonstrated that Forget-cues produce accelerated rates of forgetting compared to Remember cues and showed that these differences are independent of initial learning rates, which were deliberately equated in Experiment 2. TS cued items showed faster forgetting than Remember cued items, but were less effective than Forget cues and exhibited a more complex pattern likely reflecting individual differences. Thus, delayed testing demonstrated that active forgetting can have long-lasting effects on memory traces beyond initial suppression, in line with cognitive neuroscientific theory suggesting inhibition can produce lasting changes to memory traces.

Keywords: *directed forgetting, context shift, thought substitution, delayed testing*

Introduction

Often times, in popular culture remembering information is portrayed as the preferred outcome in memory, whereas forgetting is a failure of memory and thus an undesirable outcome. However, we regularly encounter information that is irrelevant, unimportant, or undesirable for future access, and in such cases forgetting may in fact be the desired outcome. In these cases, we may even wish to intentionally forget information in order to remove it from our mind, and a considerable amount of psychological research has been conducted to examine this process of intentional forgetting. One of the main paradigms used to investigate intentional forgetting in the lab is the Directed Forgetting (DF) paradigm (Bjork, LeBerge, & LeGrand, 1968). In studies using the DF paradigm, instructions to either remember or forget information are provided either after an entire list has been presented for study (known as the *list-method* procedure), or after the presentation of every item (known as the *item-method* procedure). Multiple studies using this method have demonstrated that participants who are instructed to forget some stimuli remember them less well than stimuli they are instructed to remember, known as the DF effect (for reviews, see Bauml, 2008; MacLeod, 1998; Sahakyan, 2023; Sahakyan & Foster, 2016; Sahakyan et al., 2013).

The theoretical mechanisms producing intentional forgetting remain a hotly debated topic, although there is a general agreement that the list-method DF effect represents a retrieval phenomenon, whereas the item-method DF effect is an encoding phenomenon. The current investigation focuses on the item-method paradigm, where the forget instruction is aimed at downregulating encoding of items in memory. According to the selective rehearsal account (Bjork, 1970; 1972; Basden, Basden, & Gargano, 1993; Hockley, Ahmad, & Nicholson, 2016; Macleod, 1975; 1999; Tan, Ensor, Hockley, Harrison, & Wilson, 2020), the DF effect emerges due to differential processing following DF cues. Namely, items are held in working memory until an instruction is presented to either remember or forget the previously presented item. A remember instruction leads to more elaborate encoding of those items, whereas a forget instruction ceases elaborate encoding, allowing items to passively decay from memory. Thus, the DF effect emerges because remember items are better encoded than forget items, rather than any active forgetting process aimed at downregulating the forget items. In contrast to the selective rehearsal account, the proponents of the inhibitory account suggests that forgetting in item-method DF arises because additional active resources are recruited to remove forget items from working memory, preventing their encoding (Fawcett & Taylor, 2008; 2010; 2012; Fawcett, Laurence, & Taylor, 2016; Festini, 2020; Hubbard & Sahakyan, 2021; 2023; Kim, Smolker, Smith, Banich, & Lewis-Peacock, 2020; Lee, 2012; Rizio & Dennis, 2013; Wylie, Foxe, & Taylor, 2008). Thus, the inhibitory account focuses on processes that downregulate forget-cued items (for review, see Anderson & Hanslmayr, 2014).

The emergence of DF effects in immediate testing has been reliably replicated across many studies over many years. However, the question of how intentional forgetting unfolds in delayed testing – that is, how the rate of intentional forgetting *itself* changes over time – is only recently gaining attention in the literature. A handful of previous studies using delayed testing procedure addressed whether DF effects persist across time in list-method designs (e.g., Abel & Bauml, 2017; 2019; Hupbach, 2018), or item-method designs (MacLeod, 1975; Schechtman et al., 2020; Simon, Gomez, and Nadel, 2018) and have largely concluded that the DF effect is obtained in delayed testing.

Delayed testing in other studies of inhibitory phenomena has produced a more mixed picture (for a review, see Anderson & Huddleston, 2012). For example, in research on retrieval-induced forgetting (RIF), some authors reported full recovery from inhibition after a day or more (Chan, 2009 ; MacLeod & Macrae, 2001 ; Saunders & MacLeod, 2002), concluding that the inhibitory effect is transient, whereas other authors have reported significant inhibition after 24 hours (Ford, Keating, & Patel, 2004; Conroy & Salmon, 2005; Conroy & Salmon, 2006; Garcia-Bajos, Migueles, & Anderson, 2009; Storm et al., 2006; Racsmány, Conway, & Demeter, 2010; Tandoh & Naka, 2007), or even after seven days (Storm, Bjork, & Bjork, 2012), suggesting more long-lasting inhibition. In think-no-think (TNT) studies, delayed testing also yielded inconsistent findings. Norby et al. (2010) found a suppression effect on an immediate test, but no suppression when those same participants were retested on the same items one week later. In contrast, Hotta and Kawaguchi (2009) found suppression on an immediate test and also a significant suppression on a re-test on those same items conducted after 24 hours.

Both Norby et al. (2010) and Hotta and Kawaguchi (2009) utilized a test-retest approach in their delayed testing, introducing complexity in the interpretation of the results. Namely, the effective use of retrieval practice enhances the subsequent retention of an item and, additionally, decelerates the rate of forgetting across extended retention periods (e.g., Karpicke & Roediger, 2008), particularly when the retrieval process is challenging. It is plausible that the initial test bolstered the items that were successfully retrieved, with varying effects depending on the difficulty of retrieval (suppressed items versus baseline items). Consequently, one could anticipate dissimilar rates of forgetting for items subjected to the initial retrieval, with baseline items being forgotten more rapidly than suppressed items. Furthermore, the initial test might have alleviated inhibition for certain items, potentially leading to an underestimation of the inhibition that could have occurred in the absence of an initial test during a delayed testing phase.

Importantly, assessment of the forgetting rates in any of these paradigms (i.e., DF, RIF, TNT) requires fitting forgetting functions to a minimum of three retention intervals because the presence or absence of these effects on immediate and delayed tests precludes conclusions about forgetting rates. Recently, Nickl and Bauml (2023) undertook such an investigation and employed an item-method DF procedure with three retention intervals, after which they fit forgetting functions to free recall and recognition test results. The authors reported that forget-cued items showed a faster forgetting rate compared to remember-cued items in both recall and recognition tests, and concluded that their findings are more consistent with the selective rehearsal account than the inhibitory account. We reexamine this issue in the current paper, and given that Nickl and Bauml (2023) is the only published study on this topic, re-addressing and extending such findings is both of theoretical and practical importance.

In contrast to Nickl & Bauml (2023), who tested different groups of participants at each delayed test, in the current investigation retention interval was manipulated within-subjects in order to directly assess forgetting over time in the *same* group of participants while avoiding repeated testing of the same items during each test. Critically, in the second experiment of our study we controlled the initial levels of learning in one of the reported experiments between the cue conditions because doing so allows addressing the theoretical debate between the selective rehearsal and inhibitory accounts more unambiguously.

In our studies, in addition to Remember and Forget cues, we also included a thought substitution (TS) cue, which involves asking participants to imagine things unrelated to the experiment. Such cues have traditionally been implemented in the list-method DF studies (Sahakyan & Kelley, 2002) and have been called "context-change cues", and have produced similar forgetting compared to the explicit instructions to forget (for a review, see Sahakyan et al., 2013). Throughout this paper, we use the terms "Imagine cues" and "TS cues" interchangeably to refer to the same process, where the mental focus or processing of the current study item is being actively replaced by thinking of another context with the help of imagination. A recent study from our lab integrated TS cues with the item-method DF design, where on some trials participants were asked to imagine specific contexts they had previously practiced imagining (Hubbard & Sahakyan, 2021). Our research indicates that TS cues are successful in inducing forgetting when applied on a trial-by-trial basis, although they appear to be less effective compared to the Forget cues in item-method DF (Hubbard & Sahakyan, 2021). Given the importance of such findings in demonstrating the contribution of contextual processes to item-method DF (see also

Chiu et al., 2021), we aimed to replicate and extend TS findings using a larger sample and compare the forgetting rates across the Forget condition and TS condition in delayed testing.

Although the majority of DF research using TS cues has been in the domain of list-method studies, where behavioral studies have shown parallel outcomes between the Forget and TS of cues (for a review, see Sahakyan et al., 2013), dissociations between Forget and TS cues have also been reported both in behavioral studies (i.e., Abel &Bauml, 2017; 2019), and electrophysiological studies (Pastotter et al., 2008). Recent studies from our own lab using the item-method DF procedure revealed electrophysiological and behavioral dissociations between Forget and TS cues (Hubbard & Sahakyan, 2021; 2023). Thought-substitution as a "strategy to forget" is frequently implemented also in TNT studies, where suppression trials (i.e., "no-think" trials) are directly contrasted with TS trials (for a review, see Anderson & Hulbert, 2021). A number of dissociations between suppression and thought substitution have been observed through electrophysiological (Bergström et al. 2009), hemodynamic (Benoit & Anderson 2012; Kim et al., 2020), and behavioral studies (Hertel & Hayes 2015, Hulbert et al. 2016, Wang et al. 2015), suggesting that perhaps unique mechanisms contribute to the occurrence of forgetting phenomena associated with these cognitive processes through these strategies. Therefore, comparison of forgetting rates between the Forget and TS cues could be informative for both theoretical and practical reasons.

Experiment 1

The purpose of Experiment 1 was to assess the rate of intentional forgetting and TS-driven forgetting across three separate retention intervals: immediately following the encoding procedure, followed by 24 hours post encoding, and again following 4 days post encoding. Based on previous research by Nickl and Bauml (2023), we expected to find a faster forgetting rate for Forget compared to Remember cued items. The TS and Forget cue conditions might exhibit parallel (accelerated) forgetting rates compared to Remember cues, or they might differ, with TS showing slower relative forgetting rate than the Forget cue. The latter prediction is based on prior research demonstrating that higher degrees of initial learning are associated with slower rates of forgetting (Wixted, 2022), and our own research indicating that in item-method designs, TS-cues lead to higher overall memory than Forget cues (Hubbard & Sahakyan, 2021). In addition, Abel and Bauml (2017; 2019) used the list-method DF design with delayed testing and included also TS cue condition. Their findings showed that on immediate test, TS and Forget cues impaired memory similarly (replicating previous research), but on a delayed test of up to 24 hours, DF effect remained significant, whereas TS-driven forgetting did not. Memory in the Forget and Remember conditions decreased with delay, whereas memory followed by TS cues did not. It could be that context shift engendered by TS cues produces only immediate impairing effects in memory, but tends to get weaker over time, which could happen if TS cues and Forget cues had different forgetting rates (i.e., faster forgetting rate in the Forget than TS condition). Testing this hypothesis requires additional retention intervals beyond immediate and delayed test to fit forgetting functions to these conditions.

Both the selective rehearsal and inhibition accounts of DF suppose that F items are learned less well than R items, and both accounts would make similar predictions for faster forgetting rates between R and F items because conditions with lower learning (i.e., F items) should be associated with faster forgetting rates (Wixted, 2022). Interestingly, Nickl and Bauml (2023) made an opposite prediction from the inhibitory account by referring to other inhibitory paradigms (i.e., RIF/TNT) and invoked the argument that inhibition could dissipate over time, which would *lower* the forgetting rate of inhibited Forget items. As mentioned previously, the dissipation of inhibition in delayed testing is far from solidly established (for a review, see Anderson & Huddleston, 2012). Most importantly, even if one assumes that retrieval inhibition in RIF/TNT studies dissipates over time, there is *a priori* reason to think that *encoding* inhibition in DF should not necessarily recover over time. If inhibitory processes are engaged to degrade the encoding of Forget items, impoverished encoding should not revert back over time.

Therefore, we had reasons to think that both the inhibitory account and the selective rehearsal account predict accelerated forgetting for F items compared to R items under conditions where initial differences in learning are observed.

Methods

Participants

A total of 146 participants participated in this online experiment and completed all delayed testing sessions (data was originally collected from 205 participants, but only 146 successfully completed all three parts of the study). All participants were recruited via the Prolific online data collection platform (www.prolific.co). They participated in April-May of 2022 and were compensated for their participation. All participants resided in the United States, spoke English as their primary language, and provided informed consent. Mean age was 31 years (range 18-44), and 91 of the participants were females. In the Hubbard and Sahakyan (2021) study, which we aimed to replicate and extend in this experiment, large effect sizes were observed for both the DF effect (Cohen's d=1.38), and thought substitution effect (Cohen's d=0.90), and a medium effect size was observed between Forget and Imagine conditions (Cohen's d=0.54). In an a priori power analysis we used the smallest effect size that we were interested in (f = 0.15) to determine the minimum required sample size, which indicated that 73 participants were required to achieve 80% power at an alpha level of .05 with repeated-measures ANOVA on a single group with three measurements (based on G*Power version 3.1.9.6, (Faul et al., 2007). We recruited more participants in anticipation of attrition across the retention intervals, and to be able to fit forgetting functions to individual participants' data given that they tend to be noisy.

Materials

Stimuli consisted of 252 nouns retrieved from the MRC Psycholinguistic Database (Coltheart, 1981). The words were medium frequency (Kucera & Francis mean word frequency of M = 50, SD = 15) and were 4-6 letters in length. Each word had an equal likelihood of being designated as old or new, and equally likely to be assigned to one of the three memory cue conditions. The memory cue conditions (Remember, Forget, and Imagine) were represented by visual images presented during the encoding phase of the study, and were taken from Hubbard & Sahakyan (2021). Remember cues were presented as a green check, while Forget cues were presented as a red X. Imagine cues consisted of a picture of a house, a utensil set, and an airplane (see Figure 1).

Design

The study employed 3 x 3 repeated-measures design, with Cue (Remember vs. Forget vs. Imagine) and Retention Interval (Day 0 vs. Day 1 vs. Day 4) varied within-subjects, with Day 0 representing immediate testing condition shortly after the study session. All participants completed three recognition tests, but on a different sample of items at each retention interval.

Procedure

Details of the procedure are shown in **Figure 1**. Prior to the DF phase, participants were provided with instructions to generate mental images for each of the three Imagine prompts, and were given 2 minutes to describe the mental image they had generated for each prompt by inputting text in a text box. As in Hubbard & Sahakyan (2021), 3 separate cues were used for the Imagine condition in order to increase the chances of participants engaging in mental contextual shifts throughout the experiment rather than repeatedly revisiting the same mental context. Text box inputs were checked following data collection, and all of the participants in the study were able to provide vivid images and details in response to the Imagine prompts.

For the Childhood Home prompt, participants received the following instruction:

> "*I want you to imagine you are in your childhood home. Imagine you are entering through the front door, and visualize the house as you travel from room to room, including details about the furniture and their location. What does it look like?*"

For the Cafeteria prompt, participants received the following instruction:

"I want you to mentally travel back in time to when you were in high school. Imagine it is time for lunch and you are walking into the cafeteria. Think of the people you sit with, the food you ate, and the smells, sounds, and layout of the room. "

For the Vacation prompt, participants received the following instruction:

"I want you to think back to a vacation you took, and picture the things you saw and activities you participated in. Where did you go? Think of how you felt, who you were with, and the experiences you had on your vacation."

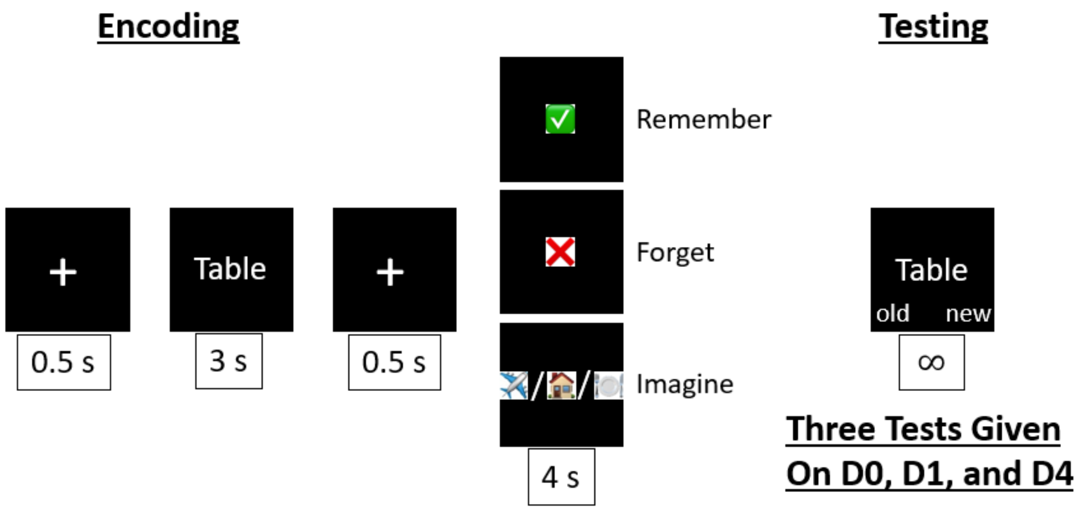


***Figure 1**. Procedure for Experiment 1. Participants studied 126 words in the DF phase, followed by an immediate Old/New recognition test (Day 0) and two delayed tests (Day 1 and Day 4). All participants were tested three times, but on a different subset of items each time. Each test consisted of a third of old words (42) intermixed with equal number of new words, for a total of 84 words in each test list. Old words were randomly selected from Remember, Forget, and Imagine conditions, and were never repeated between tests.*

After the practice with the Imagine cues, participants proceeded to the DF encoding portion of the experiment. They were instructed that they would be presented with words, each followed cues to Remember or Forget that word, or one of three Imagine cues. Specifically, participants were informed that if they received a Forget cue, those words would not be tested later and they should try to forget the word, whereas if they received a Remember cue, those words would be tested and thus they should try to remember it. Furthermore, if they received one of three Imagine cues, participants were instructed to revisit the mental image they had previously generated in the earlier Imagination instruction phase. As a cover story for the Imagine conditions, participants were informed that the purpose of the imagine cues was to investigate the effects of mind wandering on memory. No additional instructions were provided as to how participants should encode the words while they were on screen.

Following the instructions, the DF encoding phase began. On each encoding trial, a fixation point was presented for 0.5s, followed by a study word for 3s, a second fixation point for 0.5s, and finally a 4s cue presentation. The cue presentation order was randomized across trials and participants. Participants were presented with 126 study words total, with 42 words in each of the Remember, Forget, and Imagine conditions, with 14 words in each of three Imagine cue conditions (for a total of 42 overall).

Immediately following the DF encoding phase, participants were given an initial recognition test wherein they were presented with 42 old words (randomly selected from the three cue conditions) and 42 new words, one at a time, and were told to endorse each presented test word as either old or new. Importantly, at the time of test, the Forget instruction was canceled, and participants were told to endorse the word as old regardless of whether it was followed by the Forget cue. The canceling of the Forget instruction is typical in DF studies to ensure accuracy in the Forget condition is not driven by withholding information for these words. Test trials were self-paced. Participants endorsed words as either old or new by clicking on one of text boxes labeled ‘Old’ and ‘New’.

After participants completed the immediate recognition test, they were told that the second recognition test would be made available after 24 hours. The second recognition test included a different subset of 42 old (equally between Remember, Forget, and Imagine items) and 42 new words, and the instructions and procedures were identical to the initial recognition test. Following completion of the second recognition test, participants were told the third and final recognition test would be made available after 3 days (note that this means the final test was administered 4 days after the initial test). The instructions and procedure of this test were identical to the previous tests. Once the delayed test was made available, participants were given a timeframe of 36 hours to complete it, after which the study was no longer accessible. Although a 36 hour window was available to complete the delayed test, most participants completed the study shortly after it was made available online (only 5 participants took more than 24 hours to complete Day 4 testing).

Analytic Plan

Recognition accuracy (*d'*) was calculated from each participant's data in each condition after hits and false alarms were transformed using log-linear transformation to avoid estimation problems arising from hit rates of 1 and false alarm rates of 0 (Snodgrass & Corwin, 1988). Untransformed hits and false alarms are shown in Supplemental Materials, Table S1. Recognition accuracy of the initial sample (N=205) across attrition that took place at various retention intervals and cue is shown in Supplemental Figure S1.

Recognition accuracy was compared in a 3 x 3 repeated-measures ANOVA, using Cue (R vs. F vs. I) and Delay (Day 0 vs. Day 1 vs. Day 4) as factors. Significant effects of Cue were followed-up by Tukey's HSD tests to assess the DF effect and TS effect at each retention interval (i.e., the presence vs. the absence of these effects), rather than make inferences about the forgetting rates across the Cue conditions across time. ANOVA is a suitable method for analyzing this data as *d'* (unlike the data represented as proportions) has an interpretation that affords treatment on a linear scale. With that said, ANOVAs are not the suitable approach for assessing the forgetting rates because there is significant individual variability in the rates of forgetting (Kyllonen & Tirre, 1988), and averaging memory performance of participants with different forgetting rates may result in a function that fits a parametric form different from individual contributors (Anderson, 2001; Estes, 1956; Murre & Chessa, 2011). Therefore, averaged forgetting across participants assessed via ANOVAs may not accurately represent the true picture neither qualitatively nor quantitatively. Furthermore, there has been considerable debate regarding the conceptualization of forgetting, ranging from the absolute loss of units in time, to the relative loss in time (i.e., % loss when adjusted for prior baseline), with these two forgetting measures often leading to opposite conclusions, with the relative loss measure being the theoretically relevant one (Wixted, 2022). Finally, ANOVA operates under the assumption of linearity and does not take into account the retention interval between the tests. It treats the conditions as nonmetric, which is beneficial for experimental comparisons, but poorly suitable for parametric manipulations, such as manipulating the delay interval between the tests in studies on forgetting. Therefore, ANOVAs would yield identical outcomes with tests spaced at 1s, 2s, and 3s as with tests spaced at 1s, 10s, and 1000s. This poses a clear issue, which is addressed by instead fitting forgetting functions to the recognition data and estimating forgetting rates the fitted curves.

To address how different cue conditions affect the rates of forgetting, we fit recognition accuracy measures across time with a power-law forgetting function described by Wickelgren (1974): $y = a(bt+1)^c$. In this function, y represents the recognition accuracy (*d'*), *a* represents the initial degree of learning, *b* is a scaling constant, *c* represents the forgetting rate, and *t* represents time (corresponding to the delay in each testing condition, measured in days, in the current study). Forgetting functions were fit using maximum likelihood estimation (Myung, 2003). To obtain a sensible *b* parameter, we first fit a general forgetting function to averaged recognition performance of all participants collapsing across all conditions, allowing the *b* parameter to vary. The optimal *b* parameter from the general forgetting function was then used as the scaling constant for fitting individual forgetting functions, and therefore the forgetting function fit to every participant had the same *b* parameter value. To prevent extreme parameter estimation, we followed the procedure used by Siler and Benjamin (2020) and set the restrictions for the initial degree of learning (*a*) to *d'* range of 0 to 4, and forgetting rate (*c*) to a range of 0 to 1.

Estimation of forgetting parameters for individual participants is a complex task due to noisy model fits, an issue that is exacerbated if the curve does not have the canonical shape. Since our primary interest is in population-level inferences, we used the jackknife resampling procedure to derive cleaner estimates of model parameters (Efron, 1992). This method has been effectively used to derive cleaner estimates of noisy measurements such as ERP latencies (Miller et al., 1998, 2009). Here, forgetting curves were fit to N subsamples, where each subsample consisted of the averaged recognition across N-1 participants (omitting a different participant in each subsample). Thus, by using the jackknife approach, we did not fit forgetting curves to individual subjects, avoiding the issue of noisy parameter fits, but were still able to capture subject-level variability. To examine differences between cue conditions, *t*-tests were conducted on the derived initial degree of learning (*a*) and forgetting rate (*c*) parameters from the forgetting curves fit to the subsample averages across cue conditions. The subsampling procedure artificially reduces the error variances of the data; however, this can be exactly compensated for by dividing the *t*-statistics by N-1 (proof in Ulrich & Miller, 2001). Plots of the group level forgetting functions for each cue condition were based on the median parameter estimates derived across the subsamples.

Author Notes

These studies were not preregistered. Data and analysis code is available online at the following link: https://osf.io/pseqz/

Results

Recognition Accuracy. A repeated-measures ANOVA with Cue (Remember vs. Forget vs. Imagine) x Delay (D0 vs. D1 vs. D4) on recognition accuracy (*d'*) revealed a significant main effect of Cue, $F(2, 290)=34.25$, $MSE=.069$, $p<.001$, a significant main effect of Delay, $F(2, 290)=7.27$, $MSE=.212$, $p=.001$, and a significant interaction between the Cue and Delay, $F(4, 580)=2.58$, $MSE=.035$, $p=.037$ (see **Figure**

2). Next, we examined the effect of Cue at each retention interval to infer about the presence or absence of DF effect and TS effects (to enable comparisons to other published findings).

On Day 0 (immediate test), there were significant differences among the cue conditions, $F(2, 290)=27.28$, $MSE=.053$, $p<.001$. Namely, significant differences between the Remember and Forget cues revealed a DF effect, $t(145)=6.99$, $p<.001$, Cohen's $d=.58$, and significant differences between the Remember and Imagine cues revealed a TS effect, $t(145)=4.31$, $p<.001$, Cohen's $d=.36$. Finally, although both the Forget cue and Imagine cue produced forgetting compared to the Remember cue, Imagine cue produced less forgetting compared to the Forget cue, $t(145)=3.15$, $p=.002$, Cohen's $d=.26$.

On Day 1, there was a significant effect of Cue, $F(2, 290)=14.90$, $MSE=.047$, $p<.001$. Again, there were significant differences between Remember and Forget cues, $t(145)=4.84$, $p<.001$, Cohen's $d=.40$, significant differences between Remember and Imagine cues, $t(145)=3.80$, $p<.001$, Cohen's $d=.31$, and no significant differences between the Forget and Imagine cues, $t(145)=1.38$, $p=.171$, Cohen's $d=.11$.

On Day 4, there were significant differences among the cue conditions, $F(2, 290)=10.41$, $MSE=.039$, $p<.001$. There was a DF effect confirmed by significant differences between Remember and Forget cue conditions, $t(145)=4.03$, $p<.001$, Cohen's $d=.33$. There was also a significant TS effect, $t(145)=2.78$, $p=.006$, Cohen's $d=.22$. Finally, compared to the Forget cue, Imagine cues produced less forgetting, $t(145)=2.11$, $p=.037$, Cohen's $d=.18$. Thus, Day 4 findings confirmed the results of immediate test, despite overall forgetting that took place across five days in all cue conditions.

Overall, the DF effect and the TS effects were observed at each retention interval. Note that different conceptualizations on how to measure forgetting would yield opposite conclusions from this set of data if interactions in ANOVA were used as the basis of inferences. Namely, looking at the empirical data in Figure 2, it appears that both the DF effect (R-F) and the TS effects (R-I) diminished over time because recognition accuracy in the Remember condition was more adversely impacted by delay compared to the remaining cue conditions. This is evident from the absolute decline from Day 0 to Day 1 in the Remember cue (by 0.64 units) compared to Imagine (by 0.59 units) and Forget cue (by 0.50 units). The same trend is also observed from Day 1 to Day 4, with Remember cue declining by 0.16, Forget cue by 0.11, and Imagine cue by 0.09 units. Based on the absolute measure of decline, it may seem that the Remember cue has a *faster* forgetting rate compared to the remaining cue conditions. However, based on the measure of relative loss (i.e., % of loss as a function of performance at the previous retention interval), the opposite picture emerges. Namely, the Forget and Imagine cue conditions lost 63% and 62% (respectively) of their initial memory performance, compared to 53% in the Remember cue condition between Day 0 and Day 1 (the same pattern emerges also between Day 1 and Day 4). Thus, the relative measure of forgetting suggests that the forgetting rate may be *slower* (not faster) in the Remember cue condition compared to the remaining cue conditions. Since the relative rate of forgetting is the more theoretically relevant one (Wixted, 2022), and it is best estimated by fitting the forgetting functions to the data, we fit the power-law functions using the jackknife approach to examine the forgetting rates across cue conditions.

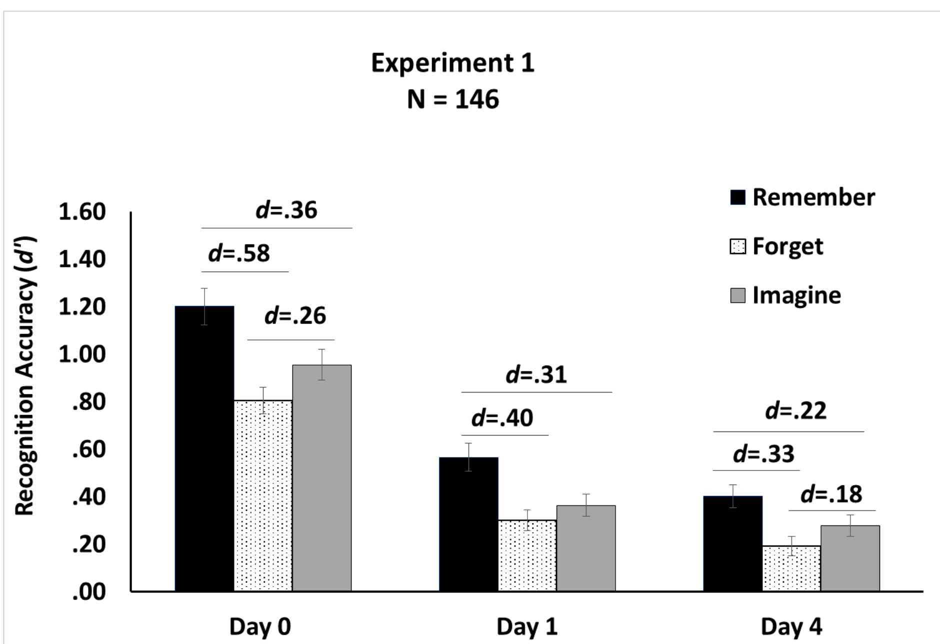


Figure 2*. Recognition Accuracy (d') across Delay and Cue in Experiment 1. Error bars represent SE of the mean.*

Estimating Forgetting Rates. For the initial degree of learning parameter (*a*), there was a significant difference between Remember (*a* = 1.20) and Forget conditions (*a* = 0.80), indicating that the Forget cue results in lower initial learning than the Remember cue does, $t(145)=7.03$, $p<.001$. We also observed a significant difference between Remember and Imagine conditions (*a* = 0.95), indicating the initial degree of learning was lower in the Imagine cue condition, $t(145)=4.37$, $p<.001$. Finally, there was a significant difference in the initial degree of learning between Imagine and Forget conditions, $t(145)=3.02$, $p=.003$. Thus, although lower initial degree of learning was observed both in Forget and Imagine cues compared to Remember cue, the magnitude of initial learning was lowest in the Forget condition. Overall, these parameter estimates are consistent with the observation of the significant DF effect and TS effects on immediate test, including the graded difference between the three cue conditions.

Critically, for the forgetting rate parameter (*c*), there was a significant difference between Remember (*c* = 0.24) and Forget (*c* = 0.31) conditions, indicating the Forget cue was associated with faster forgetting rate compared to Remember cue, $t(145)=2.04$, $p=.04$. In other words, active forgetting and passive forgetting rates seem to be different, with an intention to forget accelerating the forgetting rate of information over time. The forgetting rate for Imagine cue condition (*c* = 0.29) fell in between the Remember and Forget conditions, where it was numerically faster, but not significantly different from Remember cue ($t(145)=1.77$, $p=.08$), and also numerically slower, but not significantly different from the Forget cue, $t(145)=0.65$, $p=.51$.

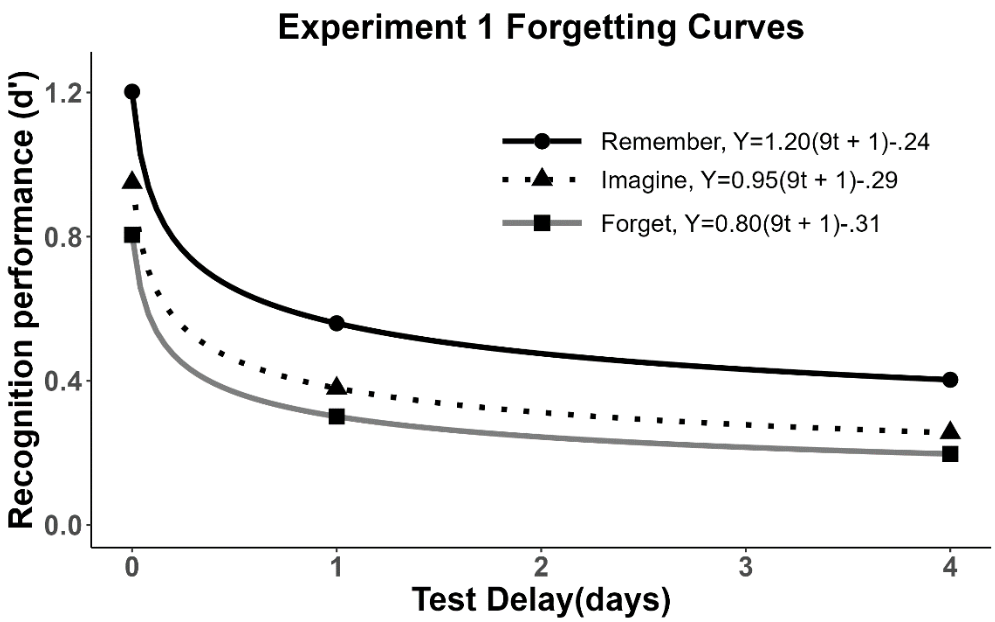


Figure 3. *Forgetting functions fit to recognition data based on median parameter estimates for each cue condition in Experiment 1.*

Discussion

Forget and TS cues resulted in impaired memory, with TS being less effective in impairing memory than Forget cues when implemented on an item-by-item basis, replicating previous work with a much larger sample (Hubbard & Sahakyan, 2021). Importantly, all of these effects were found not only on immediate test, but also in delayed tests of up to 4 days, with the decreasing accuracy across the Remember, TS, and Forget cues.

Results from the forgetting curve fitting showed that the pattern of DF and TS from initial testing was captured by parameter *a*, which reflected rates of initial learning. Namely, initial learning was highest in the Remember condition, intermediate in the Imagine cue, and lowest in the Forget cue condition. Importantly, forgetting rates captured by parameter *c* indicated that the Forget cue was associated with faster forgetting rate compared to the Remember cue, suggesting that intentional forgetting may be more accelerated than incidental forgetting in the Remember cue condition. The forgetting rate in Imagine cue condition fell in between the Remember and Forget conditions, although statistically it was not significantly different from either of them. Overall, the relative forgetting rates across the cue conditions (F > TS > R) mirrored the decreasing pattern of accuracy between R > TS > F, indicating that conditions with higher learning were associated with slower forgetting. The results of Experiment 1 support the notion that intentional forgetting via Forget cues produces greater rates of forgetting over time compared to incidental forgetting in Remember condition.

Experiment 2

In Experiment 2, we sought to equate accuracy in the immediate testing condition by familiarizing some of the items prior to subjecting them to the DF manipulation. Research from our lab indicates that familiarization significantly reduces the DF effect and may even eliminate it (Lo, Ding, Whitlock, and Sahakyan, 2023). Therefore, we expected the DF effect to be greatly reduced on initial test as a consequence of familiarization, and the critical question is whether the DF effect might emerge with delay despite being barely detectable on immediate test. Such findings would be expected based on Experiment 1 findings, indicating that Forget cue is associated with a faster forgetting rate than the Remember cue. Therefore, even if DF is not observed right away, over time it should emerge due to a faster forgetting rate in the Forget compared to the Remember condition. However, in Experiment 1, faster forgetting rates were observed in the conditions with lower initial learning (i.e., Forget and TS cues), and it is known that conditions with lower initial learning tend to be associated also with faster rate of forgetting (Wixted & Ebbesen, 1991; Wixted, 2004; 2022). Thus, differences in forgetting rates between the cue conditions in Experiment 1 can be at least partly explained by differences in their initial learning. Equating initial learning would allow for more unambiguous assessment of forgetting rates across the cue conditions.

From a theoretical perspective, both the selective rehearsal and inhibition accounts propose that F cues lead to lower levels of initial learning, and therefore should lead to faster forgetting rates, a principle demonstrated by many previous studies. However, these theories differ in their predictions for forgetting rates of equivalently learned items. If the DF effect arises solely due to differences in rehearsal rates, then equating initial learning between F and R items should yield similar forgetting rates between the R and F items. In contrast, if DF additionally arises from mechanisms that inhibit F items, then F items should nevertheless show faster forgetting even when the levels of initial learning are equated. That is, inhibition should have enduring impact on subsequent retention intervals, producing faster forgetting rates for F items.

Method

Participants

A total of 154 participants participated in this online experiment and successfully completed all three testing sessions (data was originally collected from 197 participants, but only 154 successfully completed all three parts of the study). All participants were recruited via Prolific (www.prolific.co). They participated in October-November of 2022 and were compensated for their participation. All participants resided in the United States, spoke English as their primary language, and provided informed consent. Mean age was 32 years (range 18-44), and 100 were females.

Materials

The materials were the same as in Experiment 1.

Procedure

The procedure was identical to Experiment 1, with one major exception. Prior to undergoing the DF manipulation, participants were presented with a familiarization phase in which they were shown half of the to-be-studied words (64 words) in two separate blocks in a randomized order within each block. Thus, by the end of the familiarization phase, half of the items had been seen twice. During the familiarization phase, participants were instructed to provide pleasantness judgment for each word as they were presented on screen. Afterwards, the participants were informed that they would be presented with the words they were expected to learn, some of which had been presented in the familiarization phase and some of which would be new words as they have not been seen in the familiarization phase. All further instructions and procedures followed that of Experiment 1. Each testing phase was identical to Experiment 1, and participants made the same old/new judgments regarding the test stimuli as in Experiment 1. The overwhelming majority of participants completed the delayed tests within 24 hours, with only 3 participants completing testing on Day 1 and Day 4 after 24 hours since the tests were accessible on Prolific (but before 36 hours had passed).

Design

The study employed a 3 x 3 x 2 repeated-measures design, with Cue (Remember vs. Forget vs. Imagine), Delay (Day 0 vs. Day 1 vs. Day 4) and Item type (Novel vs. Familiar).

Results

Recognition Accuracy (*d'*). The untransformed hits and false alarms are shown in Supplemental Materials, Table S2. Recognition accuracy of the initial sample (n=197) across attrition that took place at various retention intervals and cue conditions is shown in Supplemental Figure S2. Discrimination accuracy (*d'*) from the final sample that completed all three sessions (n=154) was evaluated with repeated-measures ANOVA using Cue, (R vs. F vs. I), Delay (D0 vs. D2 vs. D5), and Item type (Novel vs. Familiar) as the factors. The results are summarized in **Figure 4**.

In addition to the significant main effect of Cue, $F(2, 306)=18.66$, $MSE=.492$, $p<.001$, a significant main effect of Delay, $F(2, 306)=291.36$, $MSE=1.296$, $p<.001$, and a significant main effect of Item type, $F(1, 153)=423.55$, $MSE=1.805$, $p<.001$, there was a significant Delay x Item type interaction, *F*(2, 306)=26.05, *MSE*=.678, *p*<.001, which in turn was qualified by Cue x Delay x Item type interaction, *F*(4, 612)=3.19, *MSE*=.339, *p*=.013. None of the remaining 2-way interactions were significant (the smallest *p*=.303). To further examine the 3-way interaction, the effect of Cue and Delay were examined separately for novel and familiar items.

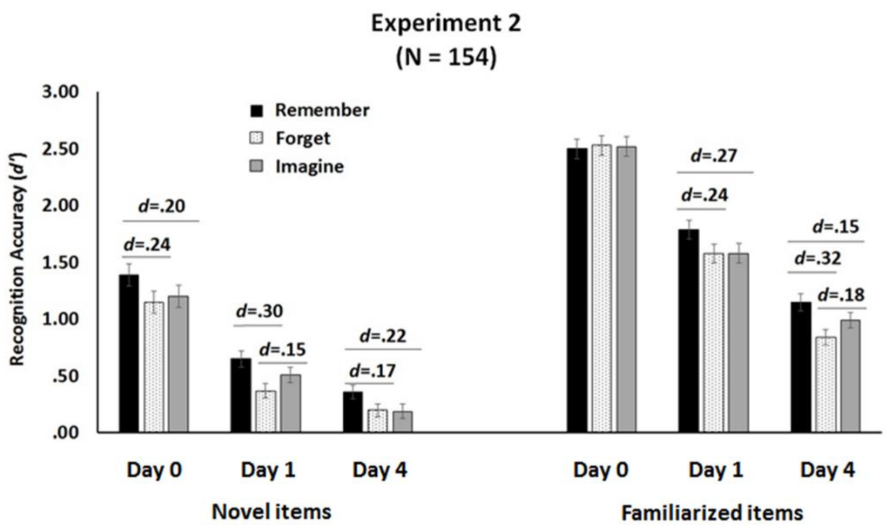


Figure 4. *Recognition Accuracy (d') across Cue, Delay, and Item type in Experiment 2. Error bars represent SE of the mean. Only effect sizes of significant differences are reported.*

For novel items, there was a significant effect of Delay, *F*(2, 306)=115.38, *MSE*=1.075, *p*<.001, denoting forgetting over time. There was also a significant main effect of Cue, *F*(2, 306)=12.93, *MSE*=.489, *p*<.001, denoting overall DF effect (*p*<.001), and overall TS effect (*p*<.001), but no difference between the Forget and Imagine cue conditions (*p*=.195). We also note that the main effect of Cue was significant in all retention intervals for novel items (Day 0: *F*(2, 306)=5.58, *MSE*=.446, *p*=.004; Day 0: *F*(2, 306)=7.01, *MSE*=.443, *p*=.001; Day 4: *F*(2, 306)=3.75, *MSE*=.362, *p*=.025). **Figure 4** left panel shows the effect sizes for all significant follow-up comparisons for novel items.

For familiarized items, in addition to the significant effect of Cue, $F(2, 306)=7.92$, $MSE=.397$, $p<.001$, and a significant effect of Delay, $F(2, 306)=301.74$, $MSE=.899$, $p<.001$, there was a significant Cue x Delay interaction, $F(4, 612)=3.86$, $MSE=.322$, $p=.004$. The interaction revealed that although there were no differences among the cue conditions on Day 0 ($F<1$), indicating that both the DF effect and the TS effects were eliminated on immediate test for familiarized items, but by Day 1, both of these effects started to emerge, as indicated by a significant effect of Cue, $F(2, 306)=6.32$, $MSE=.359$, $p=.002$. Follow-up comparisons confirmed that there were significant differences between the Remember and Forget cue conditions, $t(153)=2.99$, $p=.003$, and between the Remember and Imagine cue conditions, $t(153)=3.30$, $p=.001$, but no differences between the Forget and Imagine cues, $t<1$. **Figure 4** right panel shows the effect sizes for all significant follow-up comparisons for familiarized items.

Finally, by Day 4, significant differences between the cue conditions persisted, $F(2, 306)=7.99$, $MSE=.417$, $p<.001$. Both the DF effect and the TS effect were present five days into delayed testing, as indicated by the significant differences between the Remember and Forget cue conditions, $t(153)=4.02$, $p<.001$, and between the Remember and Imagine cue conditions, $t(153)=1.94$, $p=.050$. The difference between the Forget and Imagine cue conditions was also significant, $t(153)=2.14$, $p=.034$.

Overall, the results in immediate testing replicate our previous findings, indicating that DF is substantially reduced (and in this case, eliminated) for previously familiarized items compared to novel items. Even more important are the findings that despite initial null DF TS effects on Day 0, by Day 1 and Day 4, participants had worse memory for Forget and TS cue items than the Remember cued items, and the performance on Day 4 for familiarized items resembled Day 0 pattern for novel items, with significant DF and TS effects. This is consistent with the findings from the previous experiment, indicating that Forget and Imagine conditions are associated with accelerated forgetting rate compared to Remember condition. However, formal assessment of forgetting rates is more appropriately assessed by fitting forgetting functions.

Estimating Forgetting Rates. Results of fitting the forgetting functions to the data in the Familiarized and Novel conditions are presented in **Figure 5**. The Familiarized condition affords an opportunity to assess differences in forgetting rates that are not contaminated with differences in initial learning (Wixted & Ebbesen, 1991; Wixted, 2004; 2022), which in and of itself may have been sufficient in producing differences in forgetting rates across conditions independent of the DF manipulation.

Novel Items. When examining the initial learning rate parameter *a*, we observed significant differences between Remember (a = 1.38) and Forget (a = 1.14) conditions (t(153)=3.15, p = .002), replicating the results from Experiment 1. We also observed significant differences between Remember and Imagine (a = 1.20) conditions (t(153)=2.40, p = .02), also replicating the results from Experiment 1. However, no significant differences were observed between Forget and Imagine conditions (t(153)=0.82, p = .42). The forgetting rates were significantly different between Remember (c = 0.65) and Forget (c = 0.94), replicating the faster rate of forgetting observed for the Forget than Remember condition (t(153)=2.39, p=.02). No significant differences in forgetting rates between Remember and Imagine (c = 0.80) conditions (t(153)=1.54, p=.13), nor between the Forget and Imagine conditions (t(153)=1.11, p=.27).

Familiarized items. When examining the initial learning rate parameter *a*, we observed non-significant differences between Remember (a = 2.52) and Forget (a = 2.55) conditions (t(153)=0.41, p=.69), Remember and Imagine (a = 2.51) conditions (t(153)=0.16, p=.16), and Forget and Imagine conditions (t(153)=0.54, p=.54). These results revealed that the three cue conditions did not significantly differ in terms of initial learning, replicating our previous findings from the lab and demonstrating the effectiveness of the familiarization manipulation. Critically, despite the similar degrees of initial learning, we still observed a faster forgetting rate in the Forget cue condition (c = 0.48) compared to the Remember cue condition (c = 0.35), t(153)=4.52, p<.001. Additionally, the Imagine cue had a significantly faster forgetting rate (c = 0.42) compared to the Remember cue, $t(153)=2.91$, $p=.004$. There was a trending difference between the Forget and Imagine cue conditions, $t(153)=1.76$, $p=.08$.

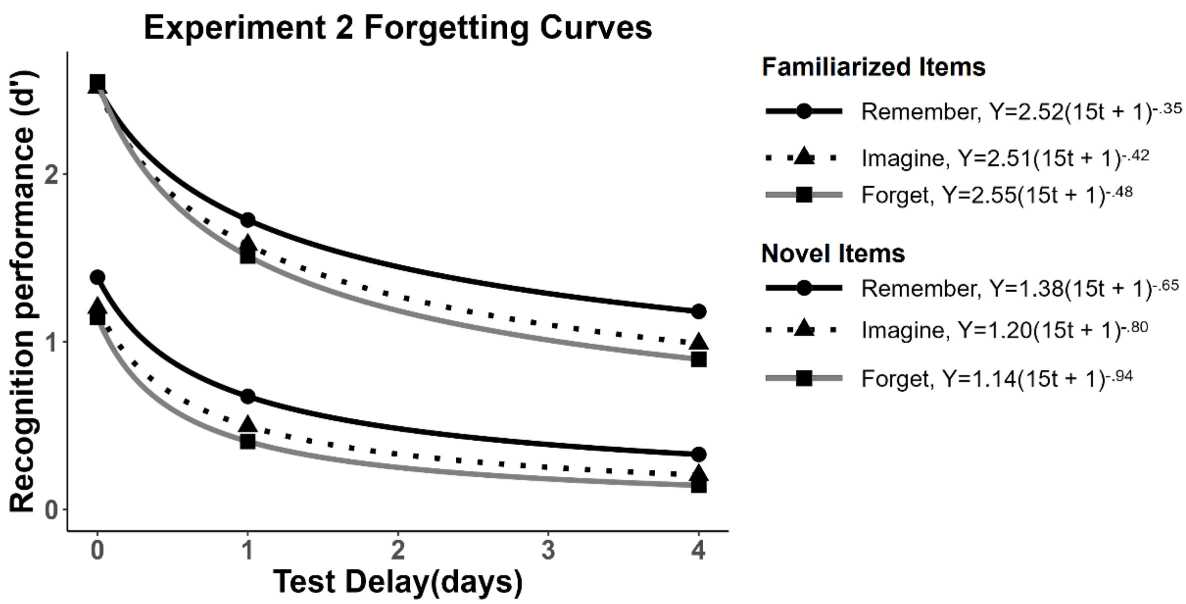


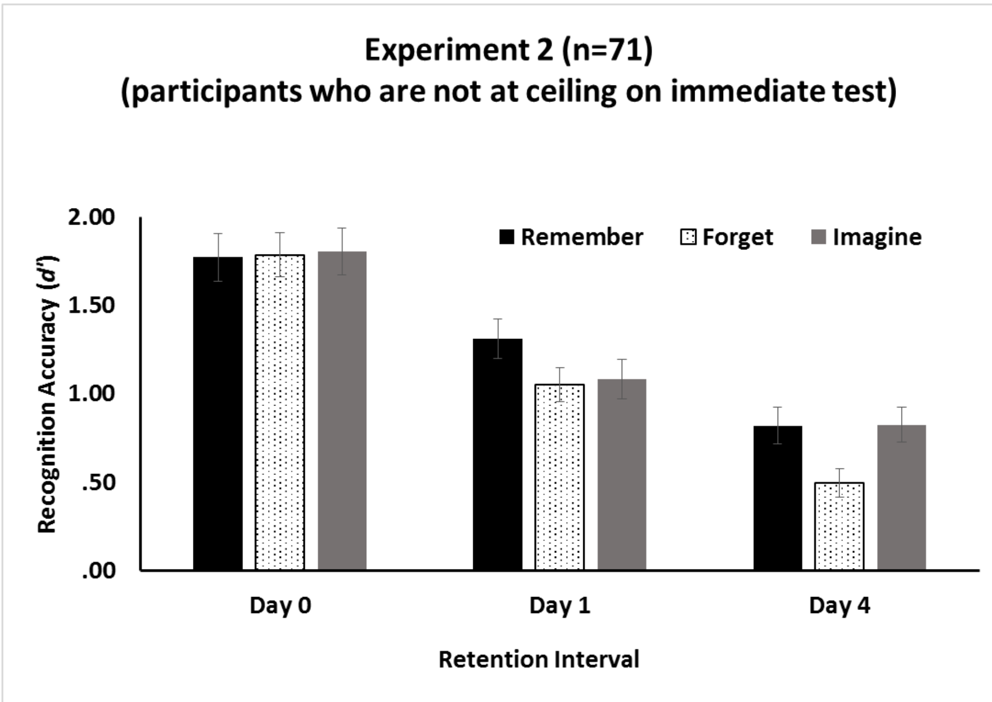
Figure 5. *Forgetting functions fit to recognition data across cue conditions and item types based on median parameter estimates in Experiment 2 (familiarized items are shown at the top portion of the figure).*

Additional Analyses of Familiarized Condition

Because the Familiarized condition is critical for adjudicating between the theories, we performed additional analyses of the data in that condition. One concern could be that despite having equated performance on the initial test, differences between the cue conditions could still exist that could be muddled by the fact that accuracy was rather high in the Familiarized condition. To address this possibility, we ran a separate analysis wherein we assessed immediate test results (Day 0) for the Familiarized items and excluded participants whose hit rate was maxed out (i.e., when adjusted hit rates $\geq$ 98%) for all cue conditions. Any remaining participants whose *d'* rate was also incredibly high (i.e., when *d'* rates were $\geq$ 3.5) were also filtered. Data from the remaining n=71 participants is visualized in **Figure 6**, top panel.

As the figure demonstrates, the overall performance on immediate test was not at ceiling, and on immediate test, there was neither a DF nor TS driven forgetting, replicating the pattern observed in the overall sample. There was, however, one notable difference – although DF emerged by Day 4 in this group of participants ($t(70)=3.37$, $p<.01$), the TS effect was absent by Day 4, ($t(70)=0.97$, $p=.34$), suggesting potentially slower forgetting rate in the Imagine cue condition. Importantly, we fit forgetting functions to the data for this subsample using the same procedures as outlined above (visualized in **Figure 7,** top panel). Consistent with previous results involving the entire sample, degrees of initial learning remained equated between Remember (a = 1.79), Forget (a = 1.80), and Imagine (a = 1.78) conditions, and yet the forgetting rates remained greater for the Forget (c = 0.56) than the Remember (c = 0.34) condition ($t(70)=3.37$, $p<.01$), as well as for the Forget than the Imagine (c = 0.40) condition ($t(70)=2.29$, $p=.03$). Forgetting rates were not significantly different between Remember and Imagine conditions ($t(70)=0.97$, $p=.33$). Thus, by excluding participants who were performing at ceiling, we nonetheless observed greater forgetting rates in the Forget than the Remember and Imagine conditions, despite equated performance on Day 0. The results are visualized in **Figure 6** (bottom panel).

We also examined DF and TS effects among the participants who were excluded from the previous analyses (i.e., high performers, n=83). **Figure 7** (top panel**),** panel shows DF and TS effects across retention interval in the "high performing" subsample. Importantly, this subsample was also equated for initial learning between Remember (a = 3.15), Forget (a = 3.19), and Imagine, (a = 3.15). Although these participants started out from much higher on Day 0 (compared to the previous subsample), by Day 4, they showed both a significant DF effect ($t(82)=2.73$, p = .008), and a significant TS effect ($t(82)=2.62$, p = .011). Also, forgetting rates were significantly faster for Forget (c = .44) than for Remember (c = .35) items ($t(82)=3.12$, p < .001), as well as faster for Imagine (c = .44) than Remember items ($t(82)=2.98$, $p < .001$). There were no significant differences between the Forget and Imagine forgetting rates, ($t(82)=0.09$, $p = .930$, indicating that TS-driven forgetting and Forget cue driven forgetting have accelerated forgetting rate among high performing participants. The results are visualized in **Figure 7** (bottom panel).



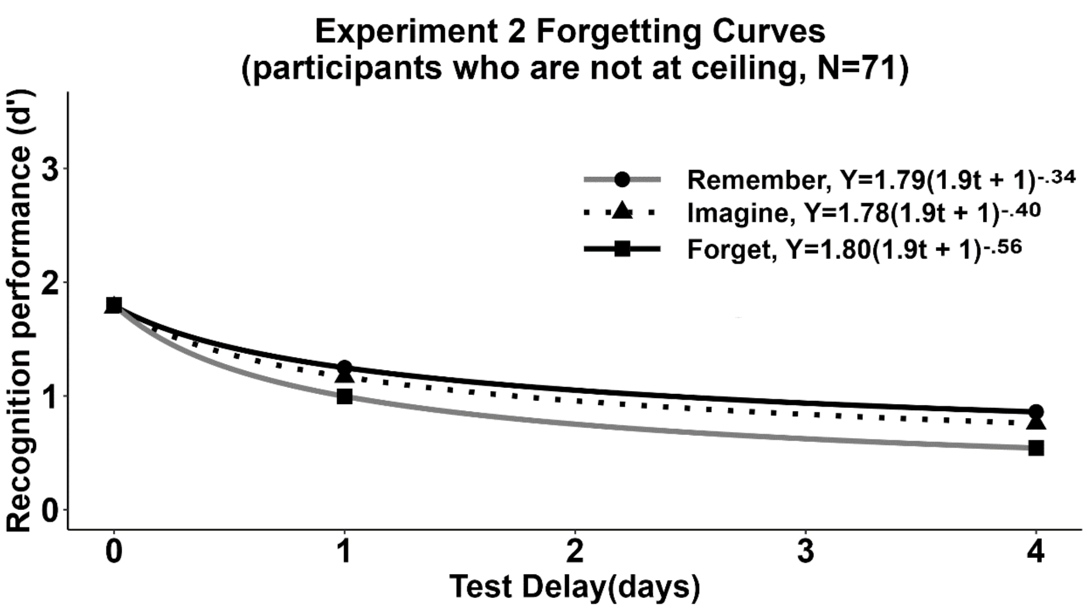
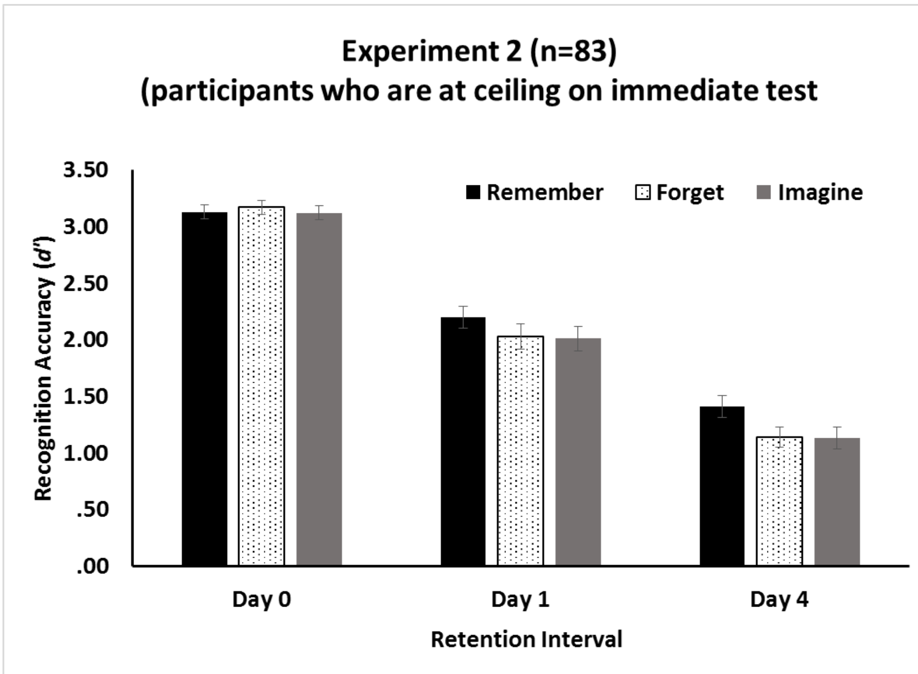


Figure 6. *Recognition Accuracy (d') for the familiarized condition across Cue and Delay in Experiment 2 for participants who were not at ceiling on immediate test. Overall accuracy is plotted in a bar graph (top panel), and the forgetting functions are shown in the bottom panel. Error bars represent SE of the mean.*



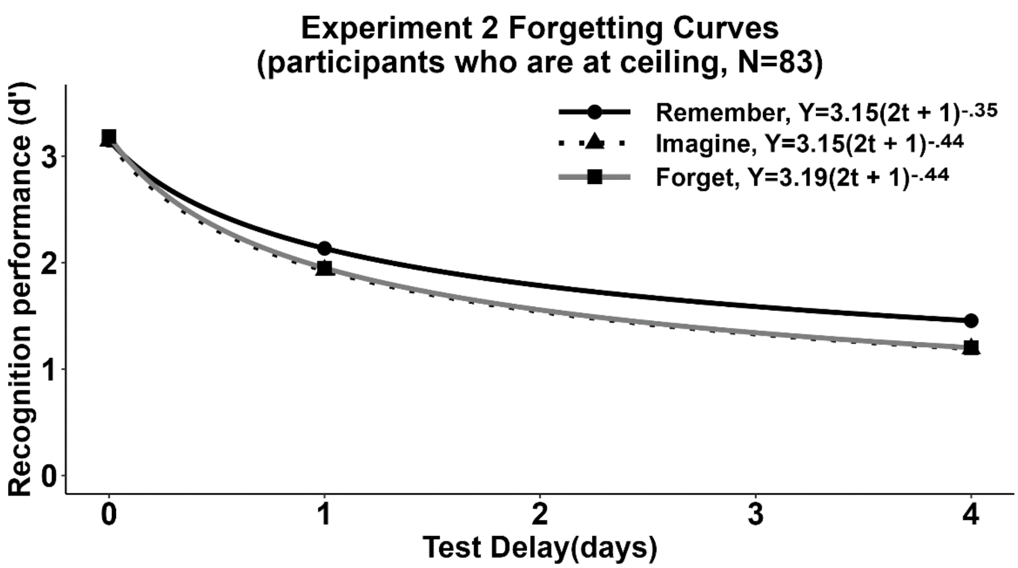


Figure 7. *Recognition Accuracy (d') for the familiarized condition across Cue and Delay in Experiment 2 for participants who were at ceiling. Overall accuracy is plotted in a bar graph (top panel), and the forgetting functions are shown in the bottom panel. Error bars represent SE of the mean.*

To summarize, regardless of individual differences reflecting whether participants were “high” or “low” performers on immediate test, we observed accelerated forgetting for F items compared to R items, despite equating for initial learning between cue conditions. Note that outcomes for TS-driven forgetting varied between high and low performers such that high performers showed accelerated forgetting both due to Imagine cues and Forget cues, unlike low performers, who had slower forgetting rate from Imagine cues compared to Forget cues. Finally, the relative rate of forgetting was similar between TS and Forget cues among high performers, but it was slower in TS than Forget cues for low performers.

Discussion

The purpose of Experiment 2 was to more unambiguously assess whether the pattern of DF and TS was due to faster forgetting rates, in which DF and TS effects would emerge across delayed testing when initial learning was equated across cue conditions. Novel items in Experiment 2 followed the general trend as observed in Experiment 1; namely, the degrees of initial learning captured the DF and TS effects, and the rate of forgetting was accelerated in the Forget compared to the Remember condition, with the Imagine condition falling between Remember and Forget conditions. When we familiarized some of the items prior to introducing the DF manipulation, we equated initial performance levels on the immediate test in the familiarized condition. Despite initial learning being similar, effects of DF and TS were evident by Day 1 and Day 4 in the familiarized condition. In fact, they resembled the pattern observed on immediate test (i.e., Day 0) of novel items. The Forget condition still produced more robust forgetting than the TS condition, showing that DF is more effective for intentional forgetting than thought substitution, even when initial memory is equivalent between the two strategies.

Critically, these findings were not contingent on high accuracy observed on immediate test in the Familiarization condition. Therefore, our results cannot easily be explained by high levels of performance masking differences between initial learning of R compared to F items. By assessing forgetting rates of non-ceiling performing participants, we demonstrated that accelerated forgetting of F items than R items continued to be observed despite equating for initial learning that was not at ceiling. We also observed this pattern in the high performing subsample. Therefore, accelerated forgetting in the Forget condition is observed regardless of the performance levels in immediate testing.

Forgetting rates in the Imagine condition, however, did show a different pattern as a function of whether participants were high or low performers. High performers showed accelerated forgetting rate due to Imagine cues and Forget cues, whereas low performers showed slower forgetting rate in Imagine cue compared to Forget cues. This outcome is consistent with prior research showing greater magnitude of TS-driven forgetting among high working memory capacity (WMC) individuals compared to low WMC (Delaney & Sahakyan, 2007); while WMC was not explicitly measured here, it is possible that the high performers in Experiment 2 were able to more effectively perform TS-driven forgetting due to a higher WMC or other individual difference measure.

General Discussion

We conducted two experiments exploring the impact of DF and TS on memory retention over time. Our aim was to test the leading accounts of DF using delayed testing as a “tool”, and to compare explicit Forget cues with Thought Substitution cues, both of which are often used to intentionally forget unwanted information. Overall, delayed testing provided insights into different forms of forgetting, shedding light on mechanisms not immediately evident. Across both experiments, immediate testing revealed intentional forgetting with both directed forgetting and thought substitution, although Forget cues were more effective than TS in an item-method DF design.

Both DF and TS effects persisted across retention intervals, showcasing their effectiveness in impairing memory consistently over time, confirming the patterns observed up to Day 4 in Experiment 1, and replicated for new items in Experiment 2. Interestingly, prior familiarization manipulation eliminated DF and TS effects on immediate testing in Experiment 2. However, despite similar initial accuracy among cue conditions, DF and TS effects emerged in delayed testing, and were especially evident by Day 4. This is important to note because if the study terminated after immediate testing, the conclusion would be that no DF effects emerge for familiarized items (or at best, that familiarization substantially reduces DF). However, the delayed testing results suggest that the intention to forget affects memory in ways that emerge over time, above and beyond that seen in immediate testing. Thought substitution also impacts memory in ways that might not be directly observable in immediate testing conditions. The emergence of DF and TS effects by Day 4 that were obscured in immediate testing suggested differences in forgetting rates.

Our findings suggest that intentional forgetting's impact unfolds over time, going beyond immediate test observations. Thought substitution also negatively affects memory in ways that might not be directly visible in immediate testing conditions. The TS cue showed impaired memory relative to the Remember cue, albeit to a lesser extent than the Forget cue. Statistical analyses using fitting of forgetting functions validated these conclusions, demonstrating faster intentional forgetting compared to incidental forgetting. Equating initial learning among cue conditions in Experiment 2 did not eliminate differences between Remember and Forget conditions, supporting the inhibitory view of intentional forgetting, indicating active processes causing accelerated forgetting over time.

In neuroscientific investigations of memory, it is well known that sensory input triggers widespread neural responses in the brain, creating unique neural patterns specific to each stimulus (Okun & Lampl, 2008; Turrigiano, 2012; Hofer, Mrsic-Flogel, Bonhoeffer, & Hubener, 2009; Xu et al., 2009). Reactivating these traces allows the retrieval of specific events and their associated context, a key aspect of episodic memory. Studies have shown that retrieving episodic memories relies on the reactivation of these specific neural representations – a phenomenon that is more prominent for remembered information compared to passively forgotten information, and it is measured through encoding-retrieval similarity (ERS) analyses.

Using intracranial EEG recordings, Ten Oever et al. (2021) found that memory traces of actively forgotten information are partially preserved and exhibit unique neural signatures. They found stronger ERS in gamma frequency for memory traces of successfully remembered R items compared to passively forgotten R items. Conversely, ERS of memory traces of actively forgotten F items depended on activity at alpha/beta frequencies commonly associated with functional inhibition, and the results were especially pronounced for successfully forgotten F items. These findings demonstrate the oscillatory processes specificity associated with active and passive forms of forgetting, and they indicate that Forget cues do not merely omit or reverse encoding-related changes but actively *modify* memory traces. These alterations occurred after initial memory formation, not erasing specific features but shaping distinct inhibitory connections. Their results indicate that while accidental forgetting reduces specific information, intentionally forgotten memories create unique representations reliant on neural patterns linked to inhibitory control. Therefore, it is fully conceivable that Forget cues might send an inhibitory "irrelevance" signal to the brain, modifying the memory traces and leading to accelerated forgetting of those traces over time. Indeed, work in our lab and others (Apšvalka et al., 2022; Hubbard & Sahakyan, 2021, 2023) suggests that domain-general inhibitory neural mechanisms are recruited when memory encoding or retrieval is actively suppressed. This rapid inhibitory process may result in a modification of the memory traces of the items themselves, though further work will be necessary to directly tie these two mechanisms together.

Interestingly, a recent article by Nickl & Bauml (2023) reported accelerated forgetting in Forget cue conditions of item-method DF, but concluded against inhibition by invoking the dissipation of retrieval inhibition over time in other inhibitory paradigms (Bauml & Kliegl, 2017; Bjork, 1989). As mentioned previously, the dissipation of inhibition in delayed testing is far from solidly established (for a review, see Anderson & Huddleston, 2012). Most importantly, effects of inhibition in item-method DF are putatively occurring at the time of encoding, whereas those observed in RIF occur in response to retrieval. Therefore, we have no reason to think that if information was inhibited at the time of learning, that information would be recovered at some later point. Importantly, the within-subjects design of our study tested separate items at each retention interval, rather than retested the same items repeatedly; thus, there was no possibility for testing to release any inhibition across testing intervals.

Even though our findings are fully consistent with an inhibitory account, and the Experiment 2 findings pose a challenge to the selective rehearsal viewpoint, accelerated forgetting in the Forget cue condition does not preclude the involvement of rehearsal processes in DF, which have been reliably documented in item-method studies. In fact, recent studies acknowledge that both inhibitory processes and rehearsal processes are contributing to item-method DF (Sahakyan & Whitlock, 2023; Fellner, Waldhauser, & Axmacher, 2020), but they have different temporal trajectories (Fellner, et al. 2020). It is likely the case that individuals engage greater rehearsal for Remember cued items, but this explanation alone is insufficient to explain the results of the current study.

Wixted (2022) proposed that the difference in forgetting rates between stronger and weaker items could be explained by the degree of learning acting as an indicator of the subjective significance of studied material. Material with higher meaningfulness might be less prone to forgetting over time, potentially due to the prioritized consolidation of this information (Cowan et al., 2021; Stickgold & Walker, 2013). Thus, If R items mimic the characteristics of highly learned items, they could benefit from prioritized consolidation, leading to a degree of protection against forgetting over time, which could additionally contribute to the observed results. Studies on the impact of sleep-associated memory consolidation on item-method directed forgetting effects support this notion (Saletin et al., 2011;

Rauchs et al., 2011), and sleep-dependent consolidation has been shown to enhance list-method directed forgetting (Blaskovich et al., 2017). However, other work suggests that items that are weakly encoded benefit more from sleep-dependent consolidation than more strongly encoded items (Bäuml, Holterman & Abel, 2014). Thus, the specific role of sleep-dependent consolidation in forgetting over time and directed forgetting needs further investigation.

One of the goals of this investigation was to compare the relative forgetting rates between explicit Forget cues and TS-cues. Whereas relative forgetting rates due to Forget cues were reliably faster than Remember cues, TS cues resulted in a more complex pattern of forgetting. In general, across two studies, the relative rates of forgetting in the TS condition fell between that in the F and R conditions, which is consistent with them producing less effective forgetting compared to Forget cues when applied on an item-by-item basis. However, there were also important exceptions. Namely, for high performers in Experiment 2, forgetting rate was faster for Imagine than Remember condition, resembling the accelerated rate of forgetting observed in the Forget condition. In contrast, low performers were less successful with TS cues, and their relative forgetting rates were slower than in Forget cues. The variability in the Imagine condition could result from incorporating multiple Imagine cues that could have varied in the extent to which they elicited mental context change, or individual differences that could have rendered specific contexts less effective than others. Given the variability of forgetting rates between high and low performers, the combination of both groups in the full sample likely contributed to the fact that in overall sample TS fell somewhere between the R and F conditions. This variance in forgetting outcomes due to TS cues is consistent with previous research comparing high and low WMC participants, with high WMC participants showing greater impairment from TS cues compared to low WMC participants (Delaney & Sahakyan, 2007). These findings have important implications for promoting active forgetting and suggest that strategies may need to be tailored to individuals who have difficulty with direct suppression. For some individuals who experience difficulty in suppressing information in memory, shifting mental context via thought substitution may be just as effective at promoting forgetting over time as direct suppression.

In conclusion, neuroscientific studies employing a concurrent neural measure which accompanies behavioral accuracy have often demonstrated dissociations between actively forgotten Forget-items (i.e., successful intentional DF) and passively forgotten Remember-items (i.e., incidental forgetting). These dissociations were reported in hemodynamic studies (Rizio & Dennis, 2012; Wylie et al., 2008), electrophysiological studies (Fellner, Waldhauser, & Axmacher, 2020; Hubbard & Sahakyan, 2021; 2023), and in eye-tracking studies (Whitlock, Lo, Chiu, & Sahakyan, 2020). In the current investigation, delayed testing played the role of a “concurrent measure” that enabled distinguishing the two forms of forgetting, providing support for inhibitory mechanisms playing a role in item-method DF. Delayed testing affords the possibility of additional scrutiny of DF mechanisms that may be obscured on immediate tests. Finally, since intentional forgetting is implemented in order to get rid of unwanted memories, it is important to study it from the perspective of how it unfolds in time as many classical studies of forgetting have done since Ebbinghaus (1885).

References

Abel, M., & Bäuml, K. H. T. (2017). Testing the context-change account of list-method directed forgetting: The role of retention interval. *Journal of Memory and Language*, *92*, 170-182.

Abel, M., & Bäuml, K. H. T. (2019). List-method directed forgetting after prolonged retention interval: Further challenges to contemporary accounts. *Journal of Memory and Language*, *106*, 18-28.

Abel, M., Haller, V., Köck, H., Pötschke, S., Heib, D., Schabus, M., et al. (2019). Sleep reduces the testing effect—But not after corrective feedback and prolonged retention interval. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 45*(2), 272-287.

Anderson, M. J. (2001). A new method for non-parametric multivariate analysis of variance. *Austral ecology*, *26*(1), 32-46.

Anderson, M. C., & Hanslmayr, S. (2014). Neural mechanisms of motivated forgetting. *Trends in cognitive sciences*, *18*(6), 279-292.

Anderson, M.C., & Huddleston, E. (2012). Towards a cognitive and neurobiological model of motivated forgetting. *Nebraska Symposium on Motivation. Nebraska Symposium on Motivation, 58*, 53-120.

Anderson, M. C., & Hulbert, J. C. (2021). Active forgetting: Adaptation of memory by prefrontal control. *Annual Review of Psychology*, 72, 1-36.

Apšvalka, D., Ferreira, C. S., Schmitz, T. W., Rowe, J. B., & Anderson, M. C. (2022). Dynamic targeting enables domain-general inhibitory control over action and thought by the prefrontal cortex. *Nature Communications*, *13*(1), 274.

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*(3), 603-616.

Bauml, K. H. T. (2008). Inhibitory processes. In H. L. Roediger, III (Ed.), *Cognitive Psychology of Memory. Learning and memory: A comprehensive reference (Vol. 2, pp. 195-220).* Oxford, Elsevier. 4 vols, (J. Byrne Editor).

Bjork, R. A., Laberge, D., & Legrand, R. (1968). The modification of short-term memory through instructions to forget. *Psychonomic Science*, *10*(2), 55-56.

Bjork, R. A. (1970). Positive forgetting: The noninterference of items intentionally forgotten. *Journal of Verbal Learning and Verbal Behavior*, *9*(3), 255-268.

Bjork, R. A. (1972). Theoretical implications of directed forgetting. In A. W. Melton & E. Martin (Eds.), Coding processes in human memory (pp. 217-325). Washing, D. C.: V. H. Winston & Sons.

Bjork, R. A. (1989). Retrieval inhibition as an adaptive mechanism in human memory. In H. L. Roediger III & F. I. M. Craik (Eds.), Varieties of memory and consciousness: Essavs in honour of Endel Tulving (pp. 309-330). Hillsdale, NJ:

Blaskovich, B., Szőllősi, Á., Gombos, F., Racsmány, M., & Simor, P. (2017). The benefit of directed forgetting persists after a daytime nap: The role of spindles and rapid eye movement sleep in the consolidation of relevant memories. *Sleep: Journal of Sleep and Sleep Disorders Research, 40*(3), 1-10.

Cowan, E. T., Schapiro, A. C., Dunsmoor, J. E., & Murty, V. P. (2021). Memory consolidation as an adaptive process. *Psychonomic Bulletin & Review*, 1-15.

Chiu, Y. C., Wang, T. H., Beck, D. M., Lewis-Peacock, J. A., & Sahakyan, L. (2021). Separation of item and context in item-method directed forgetting. *NeuroImage*, *235*, 117983.

Delaney, P. F., Sahakyan, L. (2007). Unexpected costs of high working memory capacity following directed forgetting and contextual change manipulations. *Memory & Cognition, 35*, 1074-1082.

Efron, B. (1992). *Bootstrap methods: another look at the jackknife* (pp. 569-593). Springer New York.

Estes, W. K. (1956). The problem of inference from curves based on group data. *Psychological bulletin*, *53*(2), 134.

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior research methods*, *39*(2), 175-191.

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*(6), 1168-1181.

Fawcett, J. M., & Taylor, T. L. (2010). Directed forgetting shares mechanisms with attentional withdrawal but not with stop-signal inhibition. *Memory & Cognition*, *38*(6), 797-808.

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*(1), 84-90.

Fawcett, J. M., Lawrence, M. A., & Taylor, T. L. (2016). The representational consequences of intentional forgetting: Impairments to both the probability and fidelity of long-term memory. *Journal of Experimental Psychology: General*, *145*(1), 56.

Fellner, M. C., Waldhauser, G. T., & Axmacher, N. (2020). Tracking selective rehearsal and active inhibition of memory traces in directed forgetting. *Current Biology*, *30*(13), 2638-2644.

Festini, S. B. (2020). Directed forgetting instructions decrease proactive interference within working memory below that of a baseline encode-only condition. *Journal of Cognitive Psychology*, *32*(8), 734-746.

Hockley, W. E., Ahmad, F. N., & Nicholson, R. (2016). Intentional and incidental encoding of item and associative information in the directed forgetting procedure. *Memory & Cognition, 44*, 220-228.

Hofer, S. B., Mrsic-Flogel, T. D., Bonhoeffer, T., & Hübener, M. (2009). Experience leaves a lasting structural trace in cortical circuits. *Nature*, *457*(7227), 313-317.

Hubbard, R. J., & Sahakyan, L. (2021). Separable neural mechanisms support intentional forgetting and thought substitution. *Cortex*, *142*, 317-331.

Hubbard, R. J., & Sahakyan, L. (2023). Differential recruitment of inhibitory control processes by directed forgetting and thought substitution. *Journal of Neuroscience*, *43*(11), 1963-1975.

Hupbach, A. (2018). Long-term effects of directed forgetting. *Memory*, *26*(3), 321-329.

Kim, H., Smolker, H. R., Smith, L. L., Banich, M. T., & Lewis-Peacock, J. A. (2020). Changes to information in working memory depend on distinct removal operations. *Nature Communications*, *11*(1), 6239.

Kyllonen, P. C., & Tirre, W. C. (1988). Individual differences in associative learning and forgetting. *Intelligence*, *12*(4), 393-421.

Lee, Y. S. (2012). Cognitive load hypothesis of item-method directed forgetting. *Quarterly journal of experimental psychology*, *65*(6), 1110-1122.

MacLeod, C. M. (1975). Long-term recognition and recall following directed forgetting. *Journal of Experimental Psychology: Human Learning and Memory*, *1*(3), 271.

MacLeod, C. M. (1998). Directed forgetting. In J. M. Golding & C. M. MacLeod (Eds.), *Intentional Forgetting: Interdisciplinary Approaches* (pp. 1–57). Lawrence Erlbaum Associates Publishers.

Miller, J., Patterson, T. U. I., & Ulrich, R. (1998). Jackknife-based method for measuring LRP onset latency differences. *Psychophysiology*, *35*(1), 99-115.

Miller, J., Ulrich, R., & Schwarz, W. (2009). Why jackknifing yields good latency estimates. *Psychophysiology*, *46*(2), 300-312.

Murre, J. M., & Chessa, A. G. (2011). Power laws from individual differences in learning and forgetting: mathematical analyses. *Psychonomic bulletin & review*, *18*, 592-597.

Myung, I. J. (2003). Tutorial on maximum likelihood estimation. *Journal of mathematical Psychology*, *47*(1), 90-100.

Nickl, A. T., & Bäuml, K. H. T. (2023). To-be-forgotten information shows more relative forgetting over time than to-be-remembered information. *Psychonomic Bulletin & Review*, 1-10.

Okun, M., & Lampl, I. (2008). Instantaneous correlation of excitation and inhibition during ongoing and sensory-evoked activities. *Nature neuroscience*, *11*(5), 535-537.

Pastötter B, Bäuml K-H, Hanslmayr S. 2008. Oscillatory brain activity before and after an internal context change—evidence for a reset of encoding processes. *Neuroimage, 43*, 173–81.

Rauchs, G., Feyers, D., Landeau, B., Bastin, C., Luxen, A., Maquet, P., et al. (2011). Sleep contributes to the strengthening of some memories over others, depending on hippocampal activity during learning. *The Journal of Neuroscience*, *31*(7).

Rizio, A. A., & Dennis, N. A. (2013). The neural correlates of cognitive control: successful remembering and intentional forgetting. *Journal of cognitive neuroscience*, *25*(2), 297-312.

Roediger III, H. L., Putnam, A. L., & Smith, M. A. (2011). Ten benefits of testing and their applications to educational practice. *P*

Sahakyan, L., & Kelley, C. M. (2002). A contextual change account of the directed forgetting effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*(6), 1064.

Sahakyan, L., Delaney, P. F., Abushanab*, B., & Foster*, N. L. (2013). List-Method Directed Forgetting in Cognitive and Clinical Research: A Theoretical and Methodological Review. In B. Ross (Ed.), The Psychology of Learning and Motivation, v. 59, Elsevier.

Sahakyan, L., & Foster, N. L. (2016). The need for metaforgetting: Insights from directed forgetting. *The Oxford handbook of metamemory*, *1*, 341-355.

Sahakyan, L. (2023). Current Perspectives on Directed Forgetting. In A. Wagner & M. Kahana (Eds.), *Oxford Handbook of Human Memory*.

Sahakyan, L., & Whitlock, J. (Submitted). What do people do to forget things intentionally? *Journal of Experimental Psychology: Learning, Memory, & Cognition.*

Saletin, J. M., Goldstein, A. N., & Walker, M. P. (2011). The role of sleep in directed forgetting and remembering of human memories. *Cerebral cortex*, *21*(11), 2534-2541.

Schechtman, E., Witkowski, S., Lampe, A., Wilson, B. J., & Paller, K. A. (2020). Targeted memory reactivation during sleep boosts intentional forgetting of spatial locations. *Scientific reports*, *10*(1), 2327.

Siler, J., & Benjamin, A. S. (2020). Long-term inference and memory following retrieval practice. *Memory & Cognition*, *48*, 645-654.

Simon, K. C., Gómez, R. L., & Nadel, L. (2018). Losing memories during sleep after targeted memory reactivation. *Neurobiology of Learning and Memory*, *151*, 10-17.

Snodgrass, J. G., & Corwin, J. (1988). Pragmatics of measuring recognition memory: applications to dementia and amnesia. *Journal of experimental psychology: General*, *117*(1), 34.

Tan, P., Ensor, T., Hockley, W. E., Harrison, W., & Wilson, D. (2020). In support of selective rehearsal: Double-item presentation in item-method directed forgetting. *Psychonomic Bulletin & Review, 27*(5), 529-535.

Ten Oever, S., Sack, A. T., Oehrn, C. R., & Axmacher, N. (2021). An engram of intentionally forgotten information. *Nature Communications*, *12*(1), 6443.

Turrigiano, G. (2012). Homeostatic synaptic plasticity: local and global mechanisms for stabilizing neuronal function. *Cold Spring Harbor perspectives in biology*, *4*(1), a005736.

Ulrich, R., & Miller, J. (2001). Using the jackknife-based scoring method for measuring LRP onset effects in factorial designs. *Psychophysiology*, *38*(5), 816-827.

Whitlock, J., Lo, Y., Chiu, J., & Sahakyan, L. (2020). Eye movement analyses of strong and weak memories and goal-driven forgetting. *Cognition, 204*, 10391.

Wixted, J. T., & Ebbesen, E. B. (1991). On the form of forgetting. *Psychological science*, *2*(6), 409-415.

Wixted, J. T. (2004). The psychology and neuroscience of forgetting. *Annu. Rev. Psychol.*, *55*, 235-269.

Wixted, J. T. (2022). Absolute versus relative forgetting. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 48*(12), 1775–1786.

Wylie, G. R., Foxe, J. J., & Taylor, T. L. (2008). Forgetting as an active process: An fMRI investigation of item-method–directed forgetting. *Cerebral Cortex*, *18*(3), 670-682.

Xu, T., Yu, X., Perlik, A. J., Tobin, W. F., Zweig, J. A., Tennant, K., ... & Zuo, Y. (2009). Rapid formation and selective stabilization of synapses for enduring motor memories. *Nature*, *462*(7275), 915-919.

Supplemental Materials

Table S1. *Hits and False Alarms across Cue and Delay in Experiment 1*.

		Day 0	**Day 1**	**Day 4**
		M *SD*	*M* *SD*	*M* *SD*
Hits	**Remember**	.68 (.18)	.54 (.20)	.53 (.18)
	Forget	.56 (.19)	.45 (.20)	.46 (.21)
	Imagine	.60 (.20)	.46 (.18)	.48 (.19)
False Alarms	**New**	.29 (.18)	.35 (.18)	.39 (.17)

Table S2. *Hits and False Alarms across Cue, Delay, and Familiarization in Experiment 2*

		Day 0	Day 1	Day 4
		M *SD*	*M* *SD*	*M* *SD*
Hits				
novel items	**Remember**	.61 (.30)	.48 (.26)	.46 (.26)
	Forget	.54 (.30)	.41 (.25)	.41 (.26)
	Imagine	.55 (.30)	.45 (.27)	.42 (.27)
familiarized items				
	Remember	.88 (.18)	.80 (.19)	.69 (.21)
	Forget	.88 (.18)	.74 (.23)	.60 (.25)
	Imagine	.88 (.18)	.75 (.23)	.65 (.24)
False Alarms	**New items**	.19 (.16)	.28 (.18)	.34 (.20)

Figure S1. *Recognition Accuracy (d') in Experiment 1 in Initial Sample, and Attrition across Delayed Testing and Cue. Error bars represent SE of the mean*.

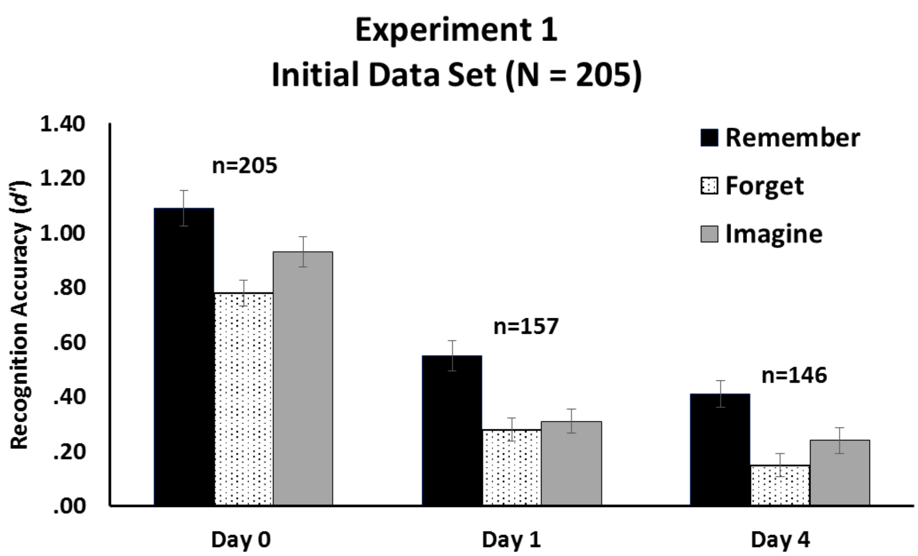


Figure S2. *Recognition Accuracy (d') in Experiment 2 in Initial Sample, and Attrition across Delayed Testing, Cue, and Familiarization. Error bars represent SE of the mean.*