AMERICAN PSYCHOLOGICAL ASSOCIATION

Journal of Experimental Psychology: General

ISSN: 0096-3445

2024, Vol. 153, No. 3, 827–836
https://doi.org/10.1037/xge0001536

Intentional Forgetting Needs Intentional Remembering

Adam Singer[1], Shira Darchi[2], Daniel Levy[3], and Talya Sadeh[1, 2, 4]
[1] Department of Cognitive and Brain Sciences, Ben-Gurion University of the Negev
[2] Department of Psychology, Ben-Gurion University of the Negev
[3] Baruch Ivcher School of Psychology, Reichman University
[4] The School of Brain Sciences and Cognition, Ben-Gurion University of the Negev

Episodic memories may become suppressed, both incidentally and intentionally. Incidental suppression is a result of a competition induced by interfering items or responses. In contrast, intentional suppression is said to result from conscious attempts to suppress certain memory items, and should thus not depend on competition induced by interfering items or responses. However, intentional suppression is typically engendered using the Think/No-Think paradigm, in which participants are required to retrieve some target items and to suppress others. Therefore, rather than intentional suppression, forgetting in this paradigm may reflect incidental suppression of No-Think items induced by interference via prior retrieval of the Think items. To distinguish between these possibilities, we tested participants ($n = 40$) using an adjusted suppression paradigm, which did not include the Think condition (ExcludeThink paradigm) and compared it with the standard suppression paradigm (IncludeThink paradigm; $n = 39$) which included a think condition. We found that suppression was not observed in the ExcludeThink paradigm, but only in the IncludeThink paradigm. These results indicate that interference via prior retrieval is necessary to induce forgetting.

Public Significance Statement

The current study questions the sweeping claim that people can weaken or even erase unwanted memories by volitional suppression alone. Rather, forgetting is induced via interference from competing items or associations. Understanding the mechanisms of induced forgetting is important because forgetting is a regulatory process that is crucial for our cognitive and affective well-being. Thus, forgetting is vital for (a) memory-regulating strategies in healthy people and (b) the development of treatments for mental disorders involving memory regulation, such as posttraumatic stress disorder and depression.

Keywords: episodic memory, forgetting, suppression, think/no-think, retrieval-induced forgetting

Supplemental materials: https://doi.org/10.1037/xge0001536.supp

It has long been acknowledged that forgetting serves important adaptive functions, which positively impact our psychological well-being and our cognitive efficiency (Bjork et al., 1998; Fawcett & Hulbert, 2020; Gamoran et al., 2020; Nørby, 2015, 2020; Sadeh & Pertzov, 2020). Forgetting is essential for emotional regulation (via selective forgetting of negative experiences), generalizing across exemplars to create categorical knowledge, and freeing up limited working-memory resources (Popov et al., 2019). Given the adaptive value of forgetting, a crucial question has concerned scientists for years: are humans able to intentionally suppress their memories?

An influential paper published in 2001 appeared to give a positive answer to this question, reporting the first evidence for suppression-induced forgetting (SIF). SIF is a putative process by which retrieval of memories can be intentionally suppressed, resulting in increased subsequent forgetting of those memories. These findings were obtained using the Think/No-Think (TNT) paradigm, which consists of three phases. First, participants learn a list of cue–

This article was published Online First January 8, 2024.
Talya Sadeh https://orcid.org/0000-0002-0717-8034
This work was supported by Israel Science Foundation, Grants 743/17 and 2055/22 to Talya Sadeh. The authors thank Matar Greenwald-Levin, May Sar Israel, Ofir Navon, and Danielle Ovadia for assistance in modifying the experiment files for the current study and in data collection. The current data and ideas have been disseminated only in internal presentations at Ben-Gurion University.

Adam Singer and Shira Darchi contributed equally to this study. Adam Singer served as lead for conceptualization and writing–original draft and contributed equally to investigation. Shira Darchi served as lead for formal analysis, validation, and visualization and served in a supporting role for writing–original draft. Daniel Levy served in a supporting role for conceptualization. Talya Sadeh served as lead for conceptualization and funding acquisition, contributed equally to writing–original draft, and served in a supporting role for methodology. Adam Singer, Daniel Levy, and Talya Sadeh contributed equally to formal analysis. Adam Singer and Daniel Levy contributed equally to methodology. Daniel Levy and Talya Sadeh contributed equally to writing–review and editing.

Correspondence concerning this article should be addressed to Talya Sadeh, Department of Cognitive and Brain Sciences, Ben-Gurion University of the Negev, David Ben Gurion Boulevard 1, 8410501 Beer Sheva, Israel. Email: tsadeh@bgu.ac.il

target word pairs (e.g., roach-ordeal). During the second phase, a third of the pairs are assigned to the Think condition, in which participants are instructed to covertly retrieve the target in response to the cue. Another third of the pairs are assigned to the No-Think (NT) condition, in which participants are instructed to suppress the target word despite the presence of the cue. The last third of the pairs, the Study-Only (SO) cues, are not presented at the second phase and serve as a control condition. In the third phase, participants perform two types of recall tests: Same-Probe (SP), in which the original cue words are used, and Independent-Probe (IP), in which novel, semantically related cue words are used. Many studies employing this paradigm have reported that the recall rate of target words from the NT condition is lower than the recall rate in the control condition, ostensibly reflecting SIF[1] (e.g., M. C. Anderson & Hanslmayr, 2014; M. C. Anderson et al., 2004; Benoit & Anderson, 2012; Detre et al., 2013; Gagnepain et al., 2014; Scotti & Maxcey, 2022) (see Stramaccia et al., 2021 for a meta-analysis). This effect is reported to generalize to other modalities of cue–target pairs, such as face–word pairs (Hanslmayr et al., 2009), and face–scene pairs (Depue et al., 2007; however, for failures to replicate see Bulevich et al., 2006; Wessel et al., 2020; Wiechert et al., 2023).

A prominent theory proposed to account for SIF is the inhibition theory. According to this theory, neural inhibitory mechanisms take part in the processes which result in forgetting (M. C. Anderson & Hanslmayr, 2014; M. C. Anderson & Huddleston, 2010; Benoit & Anderson, 2012; Gagnepain et al., 2014; Levy & Anderson, 2002). Forgetting via inhibitory mechanisms may occur either intentionally or unintentionally (Scotti & Maxcey, 2021). In TNT, neural inhibition is thought to be induced intentionally via an act of suppression, which excludes certain memories from consciousness (M. C. Anderson et al., 2004).

However, the inhibition theory of forgetting is not without opponents (Raaijmakers, 2018; Raaijmakers & Jakab, 2013; Tomlinson et al., 2009). Opposing views maintain that the forgetting effect—namely, the lower memory performance in the NT versus the SO condition—is induced by interference from competing items or responses. According to one view, the forgetting effect arises to a large extent from the vagueness of the NT instructions. This results in a substitution strategy, whereby individuals retrieve another, new word instead of the NT target during the second phase of the TNT paradigm (Raaijmakers, 2018).

An additional account also attributes forgetting in TNT to interference (Tomlinson et al., 2009). However, here, the hypothesized mechanism is different. The rationale of this account is based on two-stage global memory models of recall, which include a sampling stage, followed by a recovery stage. In the first stage, item representations are sampled from the search set, with the probability of sampling dependent on the degree of overlap between the item and contextual features. In the second stage, certain features of the sampled item become activated, and if enough of these features are activated, the item will be recovered into consciousness. Because the cue–target associations in TNT are well-learned, some NT targets are occasionally sampled automatically in the second phase of the TNT, and are associated with a so-called "sit quietly" response. Consequently, in the recovery stage during the third, and final, phase of the TNT paradigm, recall of NT items will suffer from interference from the association with the "sit quietly" response. Hence, these items are more likely to be forgotten than think and SO items. This account was supported by the finding of forgetting in a modified TNT paradigm, in which the NT condition was replaced by a condition in which individuals were instructed to do nothing in response to the cue word (Tomlinson et al., 2009). Therefore, it was concluded that, rather than intentional forgetting, the TNT effect is induced by competition from an interfering response during the third phase of final recall.

These findings raise an important question: what conditions are necessary for such interference, leading to induced forgetting? According to the interference account described above (Tomlinson et al., 2009), interference is induced in the context of a retrieval task, which involves a sampling stage. During this stage, some of the to-be-forgotten items are automatically sampled and associated with a "sit quietly" response, which interferes with these items' retrieval during final recall. Thus, interference necessitates a retrieval phase prior to the third and final recall phase, during which sampling of the to-be-forgotten items can occur. Indeed, to our knowledge, all TNT studies to date (for a meta-analysis see Stramaccia et al., 2021) included retrieval prior to final recall: the retrieving of Think items during the second phase. According to our rationale, the forgetting effect should be eliminated when excluding the think condition. This would effectively entail no retrieval phase prior to final recall, hence preventing sampling of the NT targets, and their association with an interfering response. If, on the other hand, the forgetting effect is observed even when the think condition is excluded, this would heavily shift the balance toward an intentional inhibition account.

To distinguish between these two possibilities, we implemented a modified version of the TNT paradigm in which the think condition was excluded. We based the current study on the procedure employed by Taubenfeld et al. (2019), who reported forgetting effects on explicit tests (SP and IP as described above). That study also found forgetting in an implicit test of memory, not requiring conscious retrieval and ostensibly providing a metric of forgetting not dependent on retrieval strategies. Our procedure directly followed that of Taubenfeld et al. (2019) with the exception that the think condition was excluded. To directly test whether a forgetting effect is observed only when the think condition is included, we compared results of the modified TNT paradigm excluding the think condition to those of the standard one, run in the same lab and under the same conditions as the modified TNT paradigm. The standard TNT paradigm included the think condition and was a direct replication of Taubenfeld et al. (2019). In addition, we conducted a comparison between the modified TNT paradigm and the original data from Taubenfeld et al. (2019).

Method

The method closely followed a previous study which demonstrated a TNT effect (Taubenfeld et al., 2019), with the modification that we excluded the think condition. This modified paradigm—

[1] A related phenomenon to SIF is known as list-method directed forgetting. Here, participants are instructed to forget items from a first list after their study and before learning a second list. This results in poorer recall of List 1 items compared both to List 2 items and to a condition in which participants are instructed to remember List 1 items (in this condition recall performance of List 1 exceeds that of List 2). In contrast to SIF, which involves suppression of item representations, list-method directed forgetting occurs via route deactivation. Namely, rather than inhibiting access to the items themselves as in SIF, here access to the items' retrieval routes is inhibited.

henceforth ExcludeThink—was compared to a direct replication of Taubenfeld et al. (2019), which is henceforth referred to as IncludeThink. The outline of paradigms (only depicting the SP test) is presented in Figure 1.

Participants

ExcludeThink

A total of 42 students from Ben-Gurion University of the Negev (28 female, 14 male; $M_{age} = 24.47$, $SD = 1.61$; range = 21–29) were recruited to participate in exchange for either academic credit, or payment of 80 shekels (∼$23). Informed consent was obtained from all participants according to a protocol approved by the Institutional Review Board of Ben-Gurion University of the Negev. Two participants were excluded from the analyses due to technical malfunctions in the experiment software. Participants were asked to specify their gender by selecting one of two provided options: "male" or "female." Results thus include data from 40 participants. The sample size was chosen to be identical to that in Taubenfeld et al. (2019), who obtained reliable forgetting effects.

IncludeThink

A total of 43 students from Ben-Gurion University of the Negev (30 female, 13 male; $M_{age} = 24.44$, $SD = 1.38$; range = 21–28) were recruited to participate in exchange for 12 academic credits, or payment of 120 shekels (∼$32). Informed consent was obtained from all participants according to a protocol approved by the Institutional Review Board of Ben-Gurion University of the Negev. Participants were asked to specify their gender by selecting one of two provided options: "male" or "female." Four participants were excluded from the analyses due to technical malfunctions in the experiment software. Results thus include data from 39 participants. The sample size was chosen to be identical to that of the ExcludeThink paradigm, so that any possible differences between the two paradigms could not be accounted for in terms of power.

The two groups (ExcludeThink and IncludeThink) did not differ in age, as supported by a Bayesian independent sample *t* test ($BF_{01} = 4.080$) nor in gender, as supported by a Bayesian chi-square test ($BF_{01} = 3.585$).

Materials

ExcludeThink

Following Taubenfeld et al. (2019), four exemplars were selected from 22 semantic categories (e.g., mammals: "zebra," "panther," "kangaroo," and "fox") using norms determined for Hebrew usage (Rubinsten et al., 2005). For each exemplar, based on the norms for Hebrew usage, one weakly related word was chosen as a cue. This weak semantic relation enabled the construction of a meaningful association for initial study, cued recall testing, and the Suppression phase. A second weakly related word was chosen as an independent cue, for the independent-probe recall test. For example, for "zebra," one associate was "safari" and another was "hoof," while for "fox," one associate was "forest" and another was "tail." Thus, 88 sets of triplets were created. These 88 sets were divided to three conditions: initial study test with subsequent suppression practice (33 items, NT condition); initial study test without subsequent practice (33 items, SO condition); no initial study test (22 items, category verification task test-only baseline condition). An additional 14 unrelated word pairs were constructed for use as fillers and practice trials. The assignment of triplet to condition was counterbalanced across participants.

IncludeThink

The Materials were identical to those of the ExcludeThink paradigm, with the exception that the 88 sets of triplets were divided to four conditions: initial study test with subsequent suppression practice (22 items, NT condition); initial study test with subsequent covertly retrieve practice (22 items, Think condition); initial study test without subsequent practice (22 items, SO condition); no initial study test (22 items, category verification task test-only baseline condition). Here too an additional 14 unrelated word pairs were constructed for use as fillers and practice trials.

Procedure

ExcludeThink

For an overview of the experimental procedures, see Figure 2.

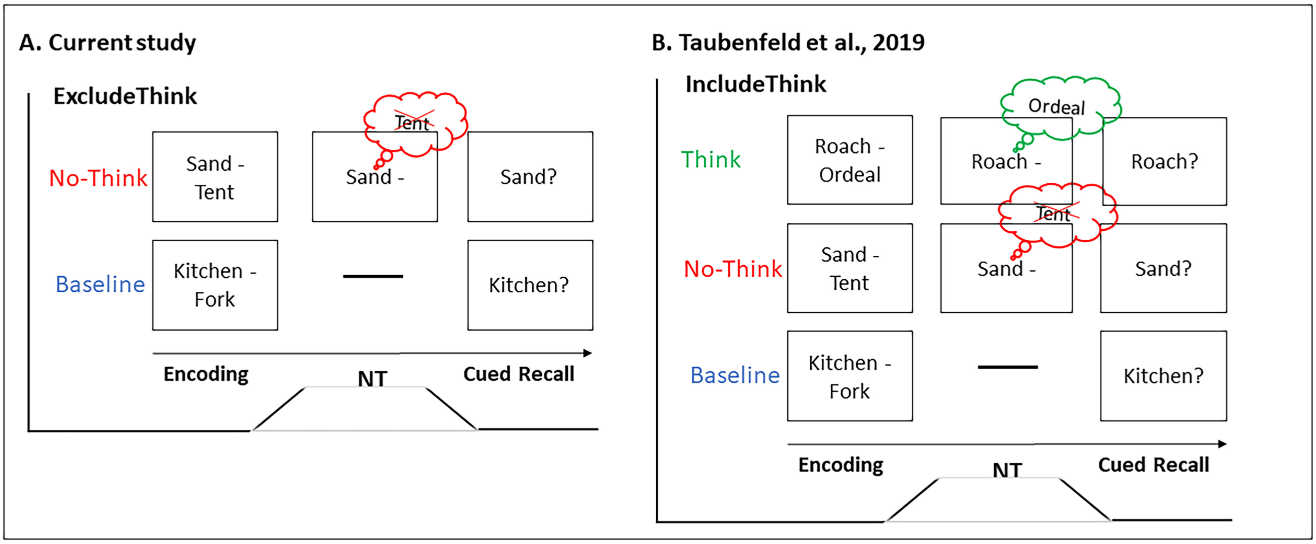
Associative Learning. In the first phase of the experiment, each participant was trained on 66 critical and 14 filler word pairs, for a total of 80 study pairs. These were divided into two lists, studied separately, each including 33 critical and seven filler word pairs, with filler pairs at the beginning and at the end of the lists and three additional fillers randomly distributed throughout the list. The word pairs were presented individually for 5 s in the center of a computer screen (black font on white background), with the cue word displayed to the right of the target (as Hebrew is read right-to-left). Participants were instructed to attempt to think of an association between the two presented words, in preparation for a later unspecified test. Trials were separated by a 1 s blank-screen interval. After learning the first list, participants were probed with cue words from that list, and asked to recall the corresponding target words and to say them aloud as quickly as possible. The correct answer was presented on screen 5 s after the cue was presented, for 2.5 s. The test phase started and ended with two fillers. The same procedure was then followed for the second list. Participants completed either two study-test cycles of both lists, or one cycle for both lists if a minimum of 60% of the targets were correctly recalled in the initial cycle.

Criterion Test. After the study-test learning phase, participants performed a criterion test, in which they were tested again on all 66 critical pairs and 14 filler word pairs from both study lists, in a random, mixed order. This test was used to determine which word pairs were successfully learned by the participants. Given that suppression can only occur in pairs which were learned in the first place, only pairs in which the participants recalled the target word successfully were considered properly learned. Accordingly, pairs in which the participants failed to recall the target word at this phase, were excluded from the analysis of later recall tests. The pairs that were included in the analysis of later recall tests constitute the conditionalized data.

Suppression (NT) Phase. After completing the study phase and the criterion test, the instructions for the suppression phase were given; participants were told that some of the studied cue words will be presented in red on the screen, and that they should try not to think of the target words which appeared alongside

Figure 1

Outline of Design of the Current Study (Only Depicting the SP Task): The ExcludeThink was a Replication of Taubenfeld et al. (2019), With the Exclusion of the Think Condition



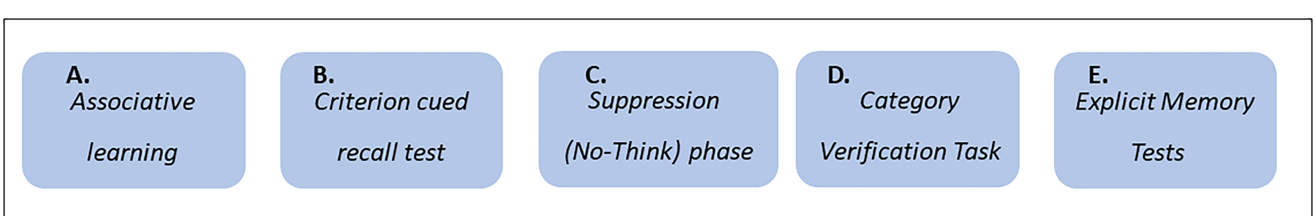
Note. The IncludeThink was a direct replication of Taubenfeld et al. (2019). SP = same-probe; NT = no-think. See the online article for the color version of this figure.

these cue words at study. Instead, when the cue appears, they should try to prevent the target word from entering their consciousness. Additionally, we employed Direct Suppression instructions (Benoit & Anderson, 2012), which stressed to participants that they were to suppress retrieval, while also not generating distracting thoughts. This started with demonstration and practice trials conducted on the 14 filler word pairs. After the practice phase, participants were asked to answer a diagnostic questionnaire to ensure that they understood the instructions and followed them closely. This procedure was then conducted for the 33 pairs, which were assigned to the NT (suppress) condition. Each of these pairs was presented 10 times across four blocks (i.e., each pair appeared two to three times per block), with each block lasting approximately 7 min. Each block started with two fillers, followed by 110 cue words from the studied pairs. Cues were presented for 3 s in the center of the screen. Trials were separated by a 500-ms blank-screen interval. After completing half of the suppression phase (e.g., after two blocks), participants were asked to answer a short questionnaire about their ability to recall or suppress responses as required, in order to maintain their level of concentration. Participants were encouraged to take a break between the blocks and at the end of the suppression phase.

Category Verification Task. After the suppression phase, subjects were given the category verification task instructions. Participants were asked to answer yes/no by key press (using the index and middle fingers of their dominant hand) to questions about category membership of target item names (e.g., "Is fig a kind of vehicle?"; "Is cider a kind of beverage?"). The target item and category names were presented in boldface. Altogether, 88 target–category pairs were presented. These pairs were divided into three conditions: 66 pairs included target words from the NT and SO conditions (33 each), while the remaining 22 pairs contained novel target words that were not presented before. The 22 pairs that contained novel target words constituted the baseline group "test-only" (TO), for studying priming effects. This task started and ended with two fillers. For the 22 categories from which four

Figure 2

Schematic Diagram of the Experimental Procedures



Note. See the online article for the color version of this figure.

words each were divided among the experimental conditions, two category members were presented in questions that were to be answered "yes" and two were presented in questions that were to be answered "no." Category names for questions that were to be answered "no" were different from those providing the exemplars for experimental conditions. Assignment of item names to yes/no responses was counterbalanced across participants. The question was displayed in the center of the screen until response, and response time (RT) was measured. Trials in which category verification decisions were incorrect were removed from response-time analyses; the mean number of such removals was 2.02, $SD = 0.73$.

Same-Probe and Independent-Probe Tests. After the category verification task, participants were given instructions for both same-probe and independent-probe tests. The same-probe test cued participants with the original cue word (e.g., for the target word "owl," the original cue word was "air"). The independent-probe test cued them with a related hint word and the first letter of the target word (e.g., "night—o_____"). In a short practice round, 14 filler word pairs were used to ensure that participants understood the instructions and did not get confused between the two tests. After this practice, participants completed the same-probe and independent-probe tests in a counterbalanced order across participants. Each test probed for all of the 66 of target words from the associative learning phase. Participants were given 5 s to recall each target word and were instructed to think of the response that fit each cue and say it aloud.

Stimuli were presented and dependent measures were collected using E-Prime 2.0 software (PST, Pittsburgh, Pennsylvania, United States). In the category verification task, trials in which participants' RT was 3 *SDs* above each participant's mean RT were removed from analyses. Statistical analyses were conducted with R Version 4.2.2 and with JASP Version 0.16.1.0. We followed up null results with Bayesian analyses, performed using R Version 4.2.2 as implemented by RStudio Version v2022.07.0+548.pro5, with the BayesFactor package (Morey et al., 2015), and with JASP.

IncludeThink

The procedure directly followed that of the ExcludeThink paradigm with the exception that the instructions for the suppression (NT) phase also included respond (think) instructions. Participants were informed that some of the studied cue words would be presented in either green or red on the screen. For green words, they were instructed to think of the target words associated with these cues.

Transparency and Openness (TOP)

Below we detail if and how the current study adheres to the eight fundamental aspects of research planning and reporting in the TOP Guidelines.

1. Citation: all data, program codes, and methods developed by others are cited in the text and listed in the references section.
2. Data transparency: the raw data are available on Open Science Framework (OSF) (https://osf.io/x4jtq/?view_only=d727347adfa44bc281aaa94778cd349d).
3. Analytic methods (code) transparency: there are no relevant scripts for data analysis.
4. Research materials transparency: all materials are available on OSF (https://osf.io/x4jtq/?view_only=d727347adfa44bc281aaa94778cd349d).
5. Design and analysis transparency: This article adheres to the APA Style Journal Article Reporting Standards (JARS; Kazak, 2018) which were deemed relevant to the current study.
6. Study preregistration: the study was not preregistered. However, the design of the ExcludeThink paradigm follows that reported in Taubenfeld et al. (2019), with the appropriate modifications to exclude the think condition. The design of the IncludeThink condition is identical to that reported in Taubenfeld et al. (2019).
7. Analysis plan preregistration: The analysis plan was not preregistered. However, the analysis of the IncludeThink condition closely follows Taubenfeld et al. (2019).
8. Replication: not relevant.

Results

ExcludeThink

Our analyses focused on the effects of retrieval suppression as expressed in both the explicit measures—recall rate in same-probe and independent-probe tests, and the implicit measure—RT in the category verification test. Since forgetting can only happen for learned items, we report analyses that consider only the conditionalized data; those items that were successfully learned in the study phase, as determined by the criterion test (mean recall success rate $= 85.8\%$, $SD = 8.3\%$). This is a frequently used procedure in TNT experiments (Hulbert et al., 2016). For a nonconditionalized analysis of all the data, including items that were not recalled in the criterion test, see the online supplemental materials.

Same-Probe and Independent-Probe Tests

For the same-probe test, a paired samples t test revealed no significant difference in recall rate between the NT ($M = 0.973$, $SD = 0.04$) and the SO ($M = 0.975$, $SD = 0.038$) conditions, $t(39) = -0.249$, $p = .804$, 95% CI [−0.015, 0.011], Cohen's $d = -0.042$ (Figure 3). Likewise, for the independent-probe test, a paired samples t test revealed no significant difference in recall rate between the NT ($M = 0.539$, $SD = 0.126$) and the SO ($M = 0.534$, $SD = 0.124$) conditions, $t(39) = 0.236$, $p = .815$, 95% CI [−0.041, 0.051], Cohen's $d = 0.043$ (Figure 3).

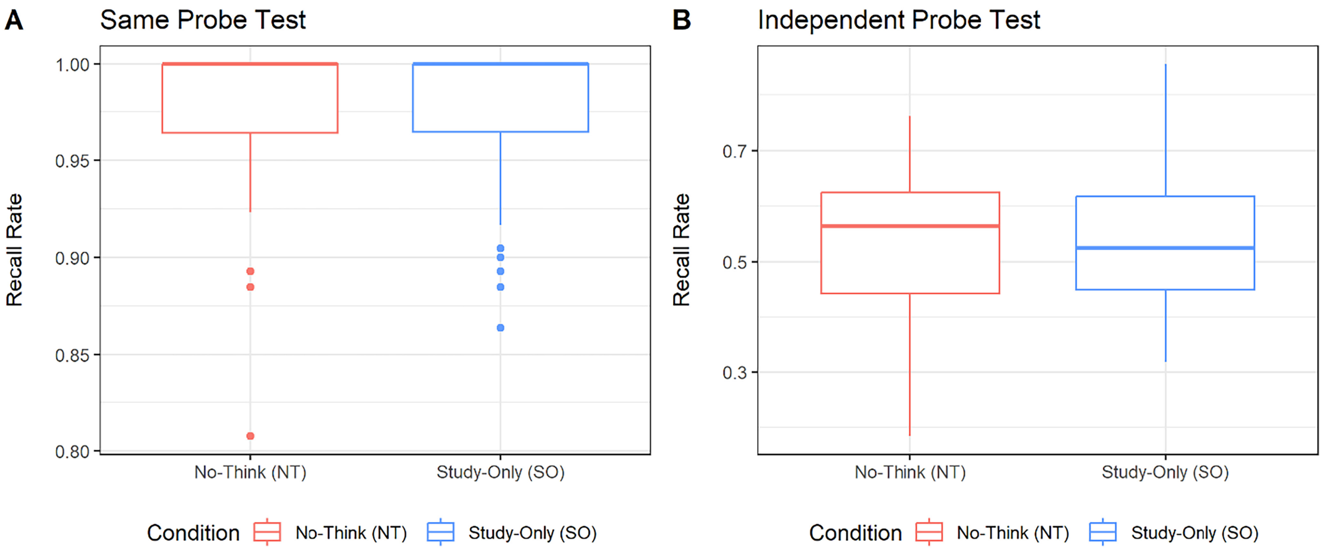
A Bayesian analysis was conducted to further examine the effects for the same-probe test and independent-probe test. Results revealed support for the null hypothesis—no differences between conditions for each task (same-probe: $BF_{01} = 5.693$; independent-probe: $BF_{01} = 5.711$).

Category Verification Task

For this analysis, we compared the RTs for words from both the NT condition and the SO condition to the RTs for words from the TO condition (Figure 4). Faster RTs were found in the NT condition ($M = 1{,}421$, $SD = 224$) and in the SO condition ($M = 1{,}448$, $SD = 264$), compared to the TO condition ($M = 1{,}532$, $SD = 263$).

A repeated measures analysis of variance (ANOVA) on the RTs with condition (NT, SO, and TO) as an independent variable was

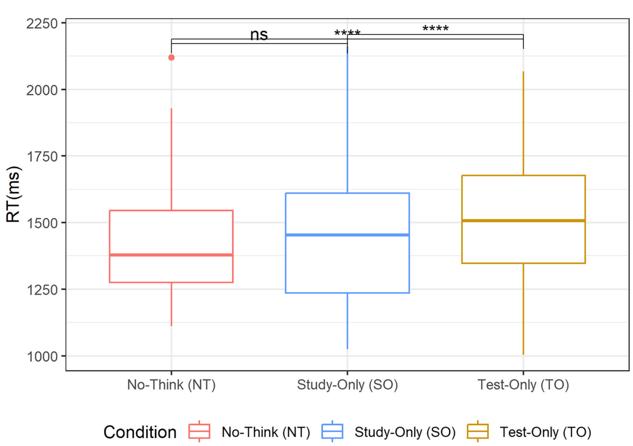
Figure 3
Results for the Same-Probe and Independent-Probe Tests



Note. Boxes depict the 25th to 75th percentiles of the data, with the median values indicated by the horizontal lines. Whiskers extend up to 1.5 times the IQR (the difference between first quartile and third quartile), and points beyond are outliers. IQR = interquartile range; SO = study-only; NT = no-think. See the online article for the color version of this figure.

conducted. Because the assumption of sphericity was violated, a Greenhouse–Geisser sphericity correction was applied. Results of the ANOVA revealed a significant effect of condition, $F(1.740, 67.843) = 15.746$, $p < .001$, $\eta_p^2 = .288$. To explore the differences between the conditions, we conducted post hoc tests. These tests showed that the difference between the NT condition ($M = 1{,}421$, $SD = 224$) and the SO condition ($M = 1{,}448$, $SD = 264$) was not significant, $t(39) = 1.313$, $p_{Holm} = .193$, 95% CI [−77.451, 23.346], Cohen's $d = 0.108$. However, significant differences were found both between the TO and the NT conditions, $t(39) = 5.382$, $p_{Holm} < .001$, 95% CI [−161.251, −60.455], Cohen's $d = 0.448$, and between the TO and the SO conditions, $t(39) = 4.068$, $p_{Holm} < .001$, 95% CI [−134, −60.455], Cohen's $d = 0.448$.

Figure 4
Results of the Category Verification Task



Note. Boxes depict the 25th to 75th percentiles of the data, with the median values indicated by the horizontal lines. Whiskers extend up to 1.5 times the IQR (the difference between first quartile and third quartile), and points beyond are outliers. IQR = interquartile range; SO = study-only; NT = no-think; TO = test-only. See the online article for the color version of this figure.

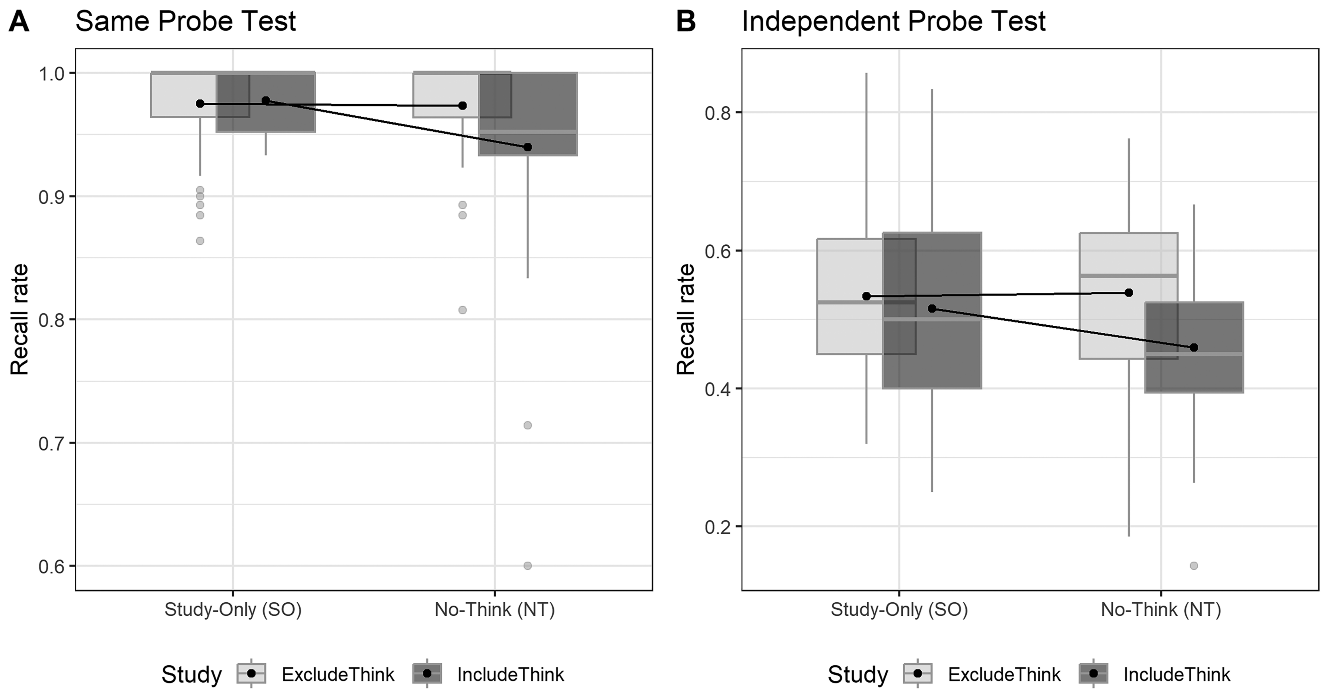
The results of a Bayesian repeated measures ANOVA on the RTs with condition (NT, SO, and TO) as an independent variable revealed strong evidence in favor of the intercept-only model over the model with the within-subjects factor of condition ($BF_{01} = 8{,}426.120$). Bayesian post hoc tests showed weak support for the null hypothesis with regard to the comparison between the NT and SO conditions ($BF_{01} = 2.947$). In contrast, there was strong support for a difference between the TO condition and both the NT condition ($BF_{10} = 1{,}095$) and the SO condition ($BF_{10} = 2{,}829$). This suggests that similar priming effects were observed for both NT and SO conditions.

Comparison With IncludeThink

The ExcludeThink paradigm was identical to that of Taubenfeld et al. (2019) in every respect except for the absence of the think condition. In the IncludeThink paradigm, we replicated the original TNT paradigm by closely following the methods of Taubenfeld et al. (2019; for full results see the online supplemental materials) and compared results of the two paradigms using a mixed ANOVA with paradigm (ExcludeThink vs. IncludeThink) as a between-subject factor and condition (SO/NT) as a within-subject factor. These analyses were aimed at examining our prediction that the inclusion (or exclusion) of the think condition should only affect performance in the NT condition. Hence, any differences between the two experiments would be specific to the NT condition, or more pronounced in it.

Figure 5

Results of the Comparison Between the ExcludeThink Paradigm (Light Gray) and IncludeThink Paradigm (Dark Gray) for the Same-Probe (A) and Independent-Probe (B) Tests



Note. Black dots denote the mean per condition. Boxes depict the 25th to 75th percentiles of the data, with the median values indicated by the horizontal lines. Whiskers extend up to 1.5 times the IQR (the difference between first quartile and third quartile), and points beyond are outliers. IQR = interquartile range; SO = study-only; NT = no-think.

Same-Probe and Independent-Probe Tests

Results of the same-probe test are displayed in Figure 5A. There was no significant main effect of paradigm, $F(1, 77) = 2.956$, $p = .090$, $\eta_p^2 = 0.037$. However, there was a significant main effect of condition, $F(1, 77) = 7.861$, $p = .006$, $\eta_p^2 = .093$, with higher recall rates in the SO condition compared to the NT condition. Most importantly, a significant interaction was found between condition and paradigm, $F(1, 77) = 6.594$, $p = .012$, $\eta_p^2 = .079$. Thus, the effects of the suppression manipulation are modulated by the inclusion/exclusion of the think condition. To further explore this interaction, we conducted post hoc tests. For the NT condition, significantly higher recall rates were found in the ExcludeThink paradigm ($M = 0.973$, $SD = 0.040$) compared to the IncludeThink paradigm ($M = 0.940$, $SD = 0.082$), $t(77) = 2.936$, $p_{\text{Holm}} = 0.015$, 95% CI [0.003, 0.065], Cohen's $d = 0.661$. For the SO condition, on the other hand, no significant difference was found between the ExcludeThink paradigm ($M = 0.975$, $SD = 0.038$) and the IncludeThink paradigm ($M = 0.978$, $SD = 0.026$), $t(77) = -0.0227$, $p_{\text{Holm}} = 1$, 95% CI [−0.033, 0.028], Cohen's $d = -0.051$. This was further supported by a Bayesian independent samples t test showing moderate support for the null hypothesis ($BF_{01} = 4.054$).

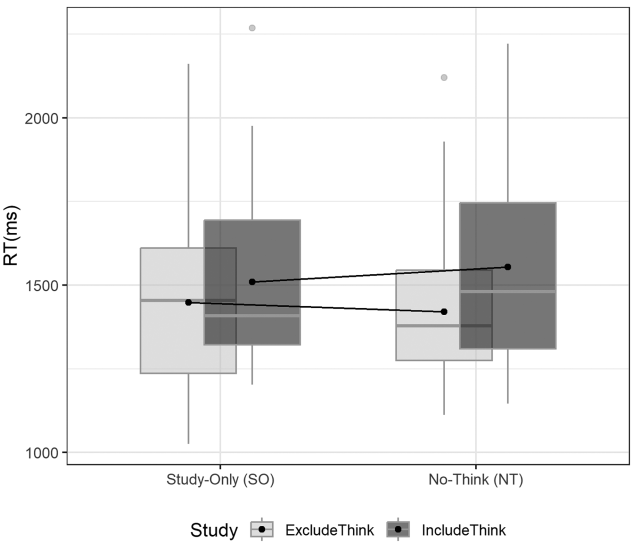
For the independent-probe test, results are presented in Figure 5B. We found no significant main effect of condition, $F(1, 77) = 2.726$, $p = .103$, $\eta_p^2 = .034$, and no significant main effect of paradigm, $F(1, 77) = 3.875$, $p = .053$, $\eta_p^2 = .048$. Most importantly, there was a significant interaction between condition and paradigm, $F(1, 77) = 4.003$, $p = .049$, $\eta_p^2 = .049$. To further explore this interaction, we conducted post hoc tests. For the NT condition, a significant difference was found between the ExcludeThink paradigm and the IncludeThink paradigm, with higher recall rates in the ExcludeThink paradigm ($M = 0.539$, $SD = 0.126$) compared to the IncludeThink paradigm ($M = 0.459$, $SD = 0.119$), $t(77) = 2.728$, $p_{\text{Holm}} = .044$, 95% CI [0.001, 0.159], Cohen's $d = 0.614$. In contrast, in the SO condition there was no significant difference between the ExcludeThink paradigm ($M = 0.534$, $SD = 0.124$) and the IncludeThink paradigm ($M = 0.516$, $SD = 0.152$), $t(77) = 0.617$, $p_{\text{Holm}} = 1$, 95% CI [−0.061, 0.097], Cohen's $d = 0.139$. This was further supported by a Bayesian independent samples t test showing moderate support for the null hypothesis ($BF_{01} = 3.693$).

Category Verification Task

Finally, we compared the results of the category verification task between the two paradigms (Figure 6). The results showed no significant main effect of paradigm, $F(1, 77) = 3.081$, $p = .083$, $\eta_p^2 = .038$, and no significant main effect of condition, $F(1, 77) = 0.247$, $p = .620$, $\eta_p^2 = .003$. However, like for the SP and IP tests, here too there was a significant interaction between condition and paradigm, $F(1, 77) = 4.146$, $p = .045$, $\eta_p^2 = .051$. This interaction shows that the differences in RTs between conditions were larger in the IncludeThink paradigm than in the ExcludeThink paradigm. Hence, the effect of the suppression manipulation differed

Figure 6

Results of the Comparison Between the ExcludeThink Paradigm (Light Gray) and the IncludeThink Paradigm (Dark Gray) for the Category Verification Task



Note. Black dots denote the mean per condition. Boxes depict the 25th to 75th percentiles of the data with the median values indicated by the horizontal lines. Whiskers extend up to 1.5 times the IQR (the difference between first quartile and third quartiles), and points beyond are outliers. IQR = interquartile range; SO = study-only; NT = no-think.

depending on the inclusion/exclusion of the think condition. To further explore this interaction, we conducted post hoc tests. However, none of the comparisons was significant. Like for the SP and IP tasks, no differences were found in RTs of the SO condition between the two paradigms ($BF_{01} = 3.070$).

The IncludeThink paradigm was an exact replication of Taubenfeld et al. (2019). Therefore, in addition to the within-experiment comparison between the ExcludeThink and the IncludeThink paradigms, we also compared data of the ExcludeThink paradigm to the data from Taubenfeld et al. (2019). Results of this comparison are detailed in the online supplemental materials. The pattern of results was the same as the within-experiment comparison reported above, with the following exceptions: (a) for the SP task, the comparison between experiments in memory scores for the NT condition only approached significance; and (b) for the category verification task, we found significantly greater priming effects in the NT condition of the ExcludeThink paradigm compared to Taubenfeld et al. (2019). Thus, this additional comparison provides further support for our conclusion that forgetting in the TNT paradigm is dependent on the inclusion of the think condition.

Discussion

In the current experiment we asked whether a forgetting effect (in the NT vs. SO conditions) occurs in the TNT paradigm even when excluding the think condition. The underlying rationale was that a forgetting effect should not be observed if it is induced by competition from an interfering response during the third phase of final recall. Most crucially, a forgetting effects should not be observed if such interference necessitates a retrieval phase prior to final recall. This retrieval phase is necessary to trigger sampling of the to-be-forgotten items and their association with an alternative "sit quietly" or "do nothing" response. When the think condition is excluded, the second phase of the TNT paradigm does not involve retrieval, hence NT items do not suffer interference via prior retrieval.

On the other hand, if suppression is intentionally induced in the NT condition, then a forgetting effect should be observed regardless of the inclusion of the think condition. A forgetting effect should also be observed according to the substitution account. In the case of substitution, individuals retrieve a new word instead of the NT target during the second phase of the TNT paradigm (Raaijmakers, 2018). Because both suppression and substitution are postulated to be a result of the NT manipulation and to occur during NT trials specifically, they should not be influenced by the existence of think trials.

The current results are most consistent with the interference via prior retrieval account. In the ExcludeThink paradigm, no significant difference was found between the recall rates of target words from the control, SO condition and target words from the NT condition, both in the same-probe and the independent-probe tests. Additionally, in the category verification test, priming effects did not differ between the NT and SO conditions. Therefore, both explicit and implicit measures indicate that no suppression occurred in the ExcludeThink paradigm.

Our results are especially striking considering that, aside from excluding the think condition, the ExcludeThink paradigm was identical to both the IncludeThink paradigm and to a previous study (Taubenfeld et al., 2019). In both, suppression effects were found, contrarily to the ExcludeThink paradigm in which suppression effects were eliminated.

Our results show that the forgetting effect necessitates interference which occurs only when there is a retrieval phase prior to final recall. This might seem similar to the retrieval-induced forgetting (RIF) phenomenon that is observed in the retrieval-practice (RP) paradigm (M. C. Anderson et al., 1994). In this paradigm, participants are first presented with a group of words from the same semantic category, which serves as their shared retrieval cue (e.g., fruits—banana, orange, apple, etc.). Then, they perform retrieval practice, in which they are required to retrieve some of the words (RP+), but not the others (RP−), in response to the retrieval cue (e.g., FRUITS-b___?). Lastly, participants preform a recall test for all the words which are associated with the retrieval cue (FRUITS - ?). RIF is expressed in results of this recall test, in that the retrieval of (RP+) items during retrieval practice (e.g., banana) impairs the recall of the (RP−) items (e.g., apple and orange), which were not practiced.

According to the Inhibition account, competition in the RP paradigm occurs between RP+ and RP− items during the retrieval practice phase. This competition results in suppression of the RP− items, hence their forgetting. In contrast, the Interference account maintains that competition results in interference during retrieval between the RP+ and the RP− items at the final test phase (for an additional, context-change account of RIF see Raaijmakers, 2018). Both these accounts of forgetting in RP differ, however, from the current account of forgetting in TNT with regard to the source of competition. In RP, the competition is from the interfering items themselves (the RP+ items). In contrast, in the current study, competition most likely arises from the association of the target with a new response.

An alternative explanation for the current findings is that the capacity for intentional suppression draws on limited resources, which do not suffice to support the on-going suppression that was required of the participants (van Schie & Anderson, 2017). However, this seems unlikely because previous studies which used a similar amount of NT items (or even larger) found suppression effects (Gagnepain et al., 2014). Hence, the current study required the same level of effort for suppressing items. If anything, task-switching (from think to NT conditions) should be more effortful than staying on the same task (J. R. Anderson & Lebiere, 2014). Even if it is the case that on-going suppression is less effective than switching between suppression and retrieval (as in the original TNT paradigm), this should require qualification of the Inhibition theory. At the very least, this elimination of the forgetting effect in the absence of a think condition, entails that suppression cannot be sustained without intervening trials which do not involve suppression.

In sum, the results of this study show that excluding the think condition from the TNT procedure eliminates the forgetting effect which is typically found in this paradigm. These results have important implications. If we want to avoid the retrieval of certain memories for the long term, interference from competing associations appears to be key. This, in turn, necessitates the presence of a retrieval phase prior to final recall. Thus, it seems that the best path to achieve forgetting is via interference from prior retrieval.

Constraints on Generality

Our reasoning regarding the generalizability of the current finding is based on findings from previous TNT paradigms which we capitalized on, all of which included a think condition. We have no reason to believe that the exclusion of the think condition in itself should affect the generalizability of our findings.

The stimuli in the current study consisted of Hebrew words. We expect our results to generalize to stimuli in different languages or even to nonverbal stimuli, as the TNT paradigm has been previously demonstrated across a wide variety of stimuli (e.g., Depue et al., 2006; van Schie et al., 2013), including naturalistic stimuli (e.g., Küpper et al., 2014; Noreen & MacLeod, 2013). Our study was run with undergraduate students serving as participants, as most previous TNT studies. We believe the results should replicate to additional populations, as previous studies have found suppression effects in the TNT paradigm with nonstudent populations (e.g., Murray et al., 2011; Sacchet et al., 2017; Waldhauser et al., 2018). We do not have evidence that the current findings will extend to settings beyond the laboratory.

References

Anderson, J. R., & Lebiere, C. J. (2014). *The atomic components of thought*. Psychology Press.

Anderson, M. C., Bjork, R. A., & Bjork, E. L. (1994). Remembering can cause forgetting: Retrieval dynamics in long-term memory. *Journal of Experimental Psychology, Learning, Memory, and Cognition*, *20*(5), 1063–1087. https://doi.org/10.1037/0278-7393.20.5.1063

Anderson, M. C., & Hanslmayr, S. (2014). Neural mechanisms of motivated forgetting. *Trends in Cognitive Sciences*, *18*(6), 279–292. https://doi.org/10.1016/j.tics.2014.03.002

Anderson, M. C., & Huddleston, E. (2010, April 22–23). *Towards a cognitive and neurobiological model of motivated forgetting* [Paper presentation]. Nebraska symposium on motivation, Nebraska (Vol. 58, pp. 53–120). https://psychology.unl.edu/symposium/recent-symposia/#sym2010

Anderson, M. C., Ochsner, K. N., Kuhl, B., Cooper, J., Robertson, E., Gabrieli, S. W., Glover, G. H., & Gabrieli, J. D. E. (2004). Neural systems underlying the suppression of unwanted memories. *Science*, *303*(5655), 232–235. https://doi.org/10.1126/science.1089504

Benoit, R. G., & Anderson, M. C. (2012). Opposing mechanisms support the voluntary forgetting of unwanted memories. *Neuron*, *76*(2), 450–460. https://doi.org/10.1016/j.neuron.2012.07.025

Bjork, E. L., Bjork, R. A., & Anderson, M. C. (1998). Varieties of goal-directed forgetting. In J. M. Golding & C. M. MacLeod (Eds.), *Intentional forgetting: Interdisciplinary approaches* (pp. 103–137). Lawrence Erlbaum Associates Publishers. https://www.researchgate.net/profile/Michael-Anderson-32/publication/259077437_Varieties_of_Goal_Directed_Forgetting/links/0046352a7278290a27000000/Varieties-of-Goal-Directed-Forgetting.pdf

Bulevich, J. B., Roediger, H. L., 3rd, Balota, D. A., & Butler, C. A. (2006). Failures to find suppression of episodic memories in the think/no-think paradigm. *Memory & Cognition*, *34*(8), 1569–1577. https://doi.org/10.3758/BF03195920

Depue, B. E., Banich, M. T., & Curran, T. (2006). Suppression of emotional and nonemotional content in memory: Effects of repetition on cognitive control. *Psychological Science*, *17*(5), 441–447. https://doi.org/10.1111/j.1467-9280.2006.01725.x

Depue, B. E., Curran, T., & Banich, M. T. (2007). Prefrontal regions orchestrate suppression of emotional memories via a two-phase process. *Science*, *317*(5835), 215–219. https://doi.org/10.1126/science.1139560

Detre, G. J., Natarajan, A., Gershman, S. J., & Norman, K. A. (2013). Moderate levels of activation lead to forgetting in the think/no-think paradigm. *Neuropsychologia*, *51*(12), 2371–2388. https://doi.org/10.1016/j.neuropsychologia.2013.02.017

Fawcett, J. M., & Hulbert, J. C. (2020). The many faces of forgetting: Toward a constructive view of forgetting in everyday life. *Journal of Applied Research in Memory and Cognition*, *9*(1), 1–18. https://doi.org/10.1016/j.jarmac.2019.11.002

Gagnepain, P., Henson, R. N., & Anderson, M. C. (2014). Suppressing unwanted memories reduces their unconscious influence via targeted cortical inhibition. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(13), E1310–E1319. https://doi.org/10.1073/pnas.1311468111

Gamoran, A., Greenwald-Levin, M., Siton, S., Halunga, D., & Sadeh, T. (2020). It's about time: Delay-dependent forgetting of item- and contextual-information. *Cognition*, *205*, Article 104437. https://doi.org/10.1016/j.cognition.2020.104437

Hanslmayr, S., Leipold, P., Pastötter, B., & Bäuml, K.-H. (2009). Anticipatory signatures of voluntary memory suppression. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *29*(9), 2742–2747. https://doi.org/10.1523/JNEUROSCI.4703-08.2009

Hulbert, J. C., Henson, R. N., & Anderson, M. C. (2016). Inducing amnesia through systemic suppression. *Nature Communications*, *7*(1), Article 11003. https://doi.org/10.1038/ncomms11003

Kazak, A. E. (2018). Editorial: Journal article reporting standards. *American Psychologist*, *73*(1), 1–2. https://doi.org/10.1037/amp0000263

Küpper, C. S., Benoit, R. G., Dalgleish, T., & Anderson, M. C. (2014). Direct suppression as a mechanism for controlling unpleasant memories in daily life. *Journal of Experimental Psychology: General*, *143*(4), 1443–1449. https://doi.org/10.1037/a0036518

Levy, B. J., & Anderson, M. C. (2002). Inhibitory processes and the control of memory retrieval. *Trends in Cognitive Sciences*, *6*(7), 299–305. https://doi.org/10.1016/S1364-6613(02)01923-X

Morey, R. D., Rouder, J. N., Jamil, T., & Morey, M. R. D. (2015). *Package "Bayesfactor."* https://cran.r-project.org/web/packages/BayesFactor/Bayes

Factor.pdf. ftp://192.218.129.11/pub/CRAN/web/packages/BayesFactor/BayesFactor.pdf

Murray, B. D., Muscatell, K. A., & Kensinger, E. A. (2011). Effects of emotion and age on performance during a think/no-think memory task. *Psychology and Aging*, *26*(4), 940–955. https://doi.org/10.1037/a0023214

Nørby, S. (2015). Why forget? On the adaptive value of memory loss. *Perspectives on Psychological Science: A Journal of the Association for Psychological Science*, *10*(5), 551–578. https://doi.org/10.1177/1745691615596787

Nørby, S. (2020). Varieties of graded forgetting. *Consciousness and Cognition*, *84*, Article 102983. https://doi.org/10.1016/j.concog.2020.102983

Noreen, S., & MacLeod, M. D. (2013). It's all in the detail: Intentional forgetting of autobiographical memories using the autobiographical think/no-think task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*(2), 375–393. https://doi.org/10.1037/a0028888

Popov, V., Marevic, I., Rummel, J., & Reder, L. M. (2019). Forgetting is a feature, not a bug: Intentionally forgetting some things helps us remember others by freeing up working memory resources. *Psychological Science*, *30*(9), 1303–1317. https://doi.org/10.1177/0956797619859531

Raaijmakers, J. G. W. (2018). Inhibition in memory. In J. Wixted (Ed.), *Stevens' handbook of experimental psychology and cognitive neuroscience Volume 1, learning and memory* (4th ed., pp. 251–284). Wiley.

Raaijmakers, J. G. W., & Jakab, E. (2013). Rethinking inhibition theory: On the problematic status of the inhibition theory for forgetting. *Journal of Memory and Language*, *68*(2), 98–122. https://doi.org/10.1016/j.jml.2012.10.002

Rubinsten, O., Anaki, D., Henik, A., Drori, S., & Faran, Y. (2005). Free association norms in the Hebrew language. In A. Henik, O. Rubinsten, & D. Anaki (Eds.), *Word norms in Hebrew* (pp. 17–34). Ben-Gurion University of the Negev.

Sacchet, M. D., Levy, B. J., Hamilton, J. P., Maksimovskiy, A., Hertel, P. T., Joormann, J., Anderson, M. C., Wagner, A. D., & Gotlib, I. H. (2017). Cognitive and neural consequences of memory suppression in major depressive disorder. *Cognitive, Affective, & Behavioral Neuroscience*, *17*(1), 77–93. https://doi.org/10.3758/s13415-016-0464-x

Sadeh, T., & Pertzov, Y. (2020). Scale-invariant characteristics of forgetting: Toward a unifying account of hippocampal forgetting across short and long timescales. *Journal of Cognitive Neuroscience*, *32*(3), 386–402. https://doi.org/10.1162/jocn_a_01491

Scotti, P. S., & Maxcey, A. M. (2021). What do laboratory-forgetting paradigms tell us about use-inspired forgetting? *Cognitive Research: Principles and Implications*, *6*(1), Article 37. https://doi.org/10.1186/s41235-021-00300-6

Scotti, P. S., & Maxcey, A. M. (2022). Directed forgetting of pictures of everyday objects. *Journal of Vision*, *22*(10), Article 8. https://doi.org/10.1167/jov.22.10.8

Stramaccia, D. F., Meyer, A.-K., Rischer, K. M., Fawcett, J., & Benoit, R. G. (2021). Memory suppression and its deficiency in psychological disorders: A focused meta-analysis. *Journal of Experimental Psychology: General*, *150*(5), 828–850. https://doi.org/10.1037/xge0000971

Taubenfeld, A., Anderson, M. C., & Levy, D. A. (2019). The impact of retrieval suppression on conceptual implicit memory. *Memory*, *27*(5), 686–697. https://doi.org/10.1080/09658211.2018.1554079

Tomlinson, T. D., Huber, D. E., Rieth, C. A., & Davelaar, E. J. (2009). An interference account of cue-independent forgetting in the no-think paradigm. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(37), 15588–15593. https://doi.org/10.1073/pnas.0813370106

van Schie, K., & Anderson, M. C. (2017). Successfully controlling intrusive memories is harder when control must be sustained. *Memory*, *25*(9), 1201–1216. https://doi.org/10.1080/09658211.2017.1282518

van Schie, K., Geraerts, E., & Anderson, M. C. (2013). Emotional and non-emotional memories are suppressible under direct suppression instructions. *Cognition & Emotion*, *27*(6), 1122–1131. https://doi.org/10.1080/02699931.2013.765387

Waldhauser, G. T., Dahl, M. J., Ruf-Leuschner, M., Müller-Bamouh, V., Schauer, M., Axmacher, N., Elbert, T., & Hanslmayr, S. (2018). The neural dynamics of deficient memory control in heavily traumatized refugees. *Scientific Reports*, *8*(1), Article 13132. https://doi.org/10.1038/s41598-018-31400-x

Wessel, I., Albers, C. J., Zandstra, A. R. E., & Heininga, V. E. (2020). A multiverse analysis of early attempts to replicate memory suppression with the think/no-think task. *Memory*, *28*(7), 870–887. https://doi.org/10.1080/09658211.2020.1797095

Wiechert, S., Loewy, L., Wessel, I., Fawcett, J. M., Ben-Shakhar, G., Pertzov, Y., & Verschuere, B. (2023). Suppression-induced forgetting: A preregistered replication of the think/no-think paradigm. *Memory*, https://doi.org/10.1080/09658211.2023.2208791

Received February 2, 2023
Revision received September 25, 2023
Accepted November 15, 2023 ■