

2018

# Ruminators (Unlike Others) Fail to Show Suppression-Induced Forgetting on Indirect Measures of Memory

Paula T. Hertel

Trinity University, [phertel@trinity.edu](mailto:phertel@trinity.edu)

Amaris Maydon

Trinity University, [amaydon@trinity.edu](mailto:amaydon@trinity.edu)

Ashley Ogilvie

Trinity University, [aogilvie@trinity.edu](mailto:aogilvie@trinity.edu)

Nilly Mor

Follow this and additional works at: [https://digitalcommons.trinity.edu/psych\\_faculty](https://digitalcommons.trinity.edu/psych_faculty)

 Part of the [Psychology Commons](#)

Publication Details

Clinical Psychological Science

---

## Repository Citation

Hertel, P. T., Maydon, A., Ogilvie, A., & Mor, N. (2018). Ruminators (unlike others) fail to show suppression-induced forgetting on indirect measures of memory. *Clinical Psychological Science*. Advance online publication. doi:10.1177/2167702618785339

This Post-Print is brought to you for free and open access by the Psychology Department at Digital Commons @ Trinity. It has been accepted for inclusion in Psychology Faculty Research by an authorized administrator of Digital Commons @ Trinity. For more information, please contact [jcostanz@trinity.edu](mailto:jcostanz@trinity.edu).

Ruminators (Unlike Others) Fail to Show Suppression-Induced Forgetting  
on Indirect Measures of Memory

Paula T. Hertel, Amaris Maydon, Ashley Ogilvie

Trinity University

Nilly Mor

Hebrew University

Author Note

Paula T. Hertel, Amaris Maydon, and Ashley Ogilvie, Department of Psychology, Trinity University; Nilly Mor, School of Education, Hebrew University, Israel

This research was funded by a United States/Israel Binational Science Foundation Grant to the first and last authors (BSF 2011267).

Correspondence concerning this article should be addressed to Paula Hertel, Department of Psychology, Trinity University, San Antonio, TX 78212; phertel@trinity.edu.

## Abstract

Suppression is a useful everyday skill leading to the clinically important outcome of forgetting. Suppression-induced forgetting, investigated with the think/no-think (TNT) paradigm, is typically demonstrated on direct tests of memory, even though indirect tests are often more ecologically valid. We report results from two TNT experiments terminating in indirect tests—tests that *seem* not to measure memory. For a subset of the participants in Experiment 1, latencies to rate word valence were delayed by flankers previously learned but not by flankers previously learned and then suppressed on 16 occasions. For a similar subset in Experiment 2, cue meaning denoted by free associations reflected the meaning established during learning, but less so when targets had been suppressed. These subsets showing suppression effects were students who did not describe themselves as ruminators. That ruminators failed to show “forgetting” on either test suggests that ruminative habits cannot be easily overcome by practicing suppression.

[150 words]

Key words: forgetting, suppression, rumination, habit, depression

Ruminators (Unlike Others) Fail to Show Suppression-Induced Forgetting  
on Indirect Measures of Memory

The term *forget* is typically used to refer to the unsuccessful outcome of a deliberate attempt to remember, but in this report we address a different and perhaps more pervasive type of forgetting in which past experience fails to influence current behavior, as it might otherwise without being deliberately sought (see Bjork, 1989). By definition, this form of forgetting often goes unnoticed, but it is nonetheless valuable in a host of ways, such as when reminders initiate maladaptive obsessions or ruminations, replete with intentional recall. In this sense, forgetting “without awareness” can be therapeutically important and therefore might be worthy of efforts to develop procedures to encourage the use of suppression. And so, following a long line of research on suppression-induced forgetting on direct memory tests with the procedure called the think/no-think paradigm (TNT; see the review by Anderson & Huddleston, 2012), we report the outcomes from two experiments designed to reveal forgetting in the form of transfer of suppression across apparently unrelated tasks.

TNT experiments consist of three phases: (a) an initial phase for learning cue-response pairs; (b) a TNT phase in which some cues repeatedly invite recall of their responses (think), other cues repeatedly signal suppression (no-think), and still others are not presented (baseline); and (c) typically a final direct test of cued recall of all responses. As expected from a transfer-appropriate-processing perspective, practicing recall in the TNT phase produces above-baseline recall on the test, and practicing suppression produces below-baseline recall. From an ecological perspective, however, direct tests of suppression-induced forgetting are inferior to indirect tests, because people who expend great efforts trying not to think about negative events would rarely

attempt their deliberate recall later on (unless prompted by clinical or legal professionals).<sup>1</sup> In this first respect, then, our experiments were designed to extend evidence already accumulating that experience with TNT procedures transfers in ecologically interesting ways to tasks that are capable of either indirectly reminding or reflecting prior suppression (see the review by Hue, Bergstrom, Gagnepain, & Anderson, 2017). As is the case for direct tests, the difference between performance on baseline and suppressed items serves as the indirect measure of forgetting.

Indirect tests, if they are designed well, are interesting measures of the extent to which suppression attempts can truly change cognitive habits. Direct testing might inadvertently suggest to the TNT participant that it is time to do what has been done most recently—suppress or respond—whereas indirect tests can index a broader base of transfer from the situations in which memory cues might have initially been encountered. This ecological advantage for indirect tests is fundamentally related to our interest in ruminative thinking.

People who report that, when they are sad, they habitually ask themselves certain questions (e.g., “Why do I always react this way?”) or think certain thoughts (e.g., about their shortcomings, faults, mistakes) earn high scores on the Ruminative Response Scale of the Response Styles Questionnaire (RRS; Treynor, Gonzales, & Nolen-Hoeksema, 2003) or similar self-report instruments. Defined in this way, ruminators perform in experimental tasks in ways that indicate deficient cognitive control, particularly when the context is flexible (e.g., Joormann, Levens, & Gotlib, 2009; see Yang, Cao, Shields, Teng, & Liu, 2017). Yet, regardless of the *ability* to exert control, doing so for the purpose of suppressing certain thoughts or memories should indeed be difficult for those who fail to practice such control routinely (Hertel, 2004; Watkins & Nolen-Hoeksema, 2014). Ruminators practice the opposite; they repeatedly attend to

---

<sup>1</sup> Justin Hulbert (a reviewer) suggested an exception in cases in which we expend great efforts to repeatedly postpone thinking about a distraction that we subsequently want to address when the current task is finished.

and deliberately recall matters of personal concern in a variety of contexts. Ruminative thoughts can be invoked automatically, but once invoked they are practiced with intention. Due to habits of ruminating in a variety of contexts, suppression practice should not only be difficult to do well but (like extinction or counterconditioning trials, Bouton, 2000; Miller & LaBorda, 2011) should also be difficult to generalize to new situations, which is what is required from well-designed indirect tests. In the following experiments, we therefore investigated suppression-induced forgetting (SIF) on two indirect tests and predicted rumination-related impairments on both. Finding such evidence would reveal that the retrieval habits associated with depressive rumination maintain the availability of ruminative thoughts by making suppression difficult.

### **Experiment 1**

The first experiment is a variation of experiments that showed evidence of SIF on a flanker task (Hertel & Hayes, 2015). In those previous experiments, critical flanker trials each consisted of a display of a central, novel word to be rated for emotional valence, surrounded above and below by a flanker word that varied according to its role in the prior TNT phase. The flanker was either new itself, or it had previously cued a baseline or suppressed response word from Phases 1 and 2 of the TNT paradigm. Ratings for the central word were slowed by baseline flankers (a cue to think a particular thought), compared to both new and suppression flankers. In the present experiment, response words instead of cues served as flankers. We reasoned that, because distraction by baseline (compared to new) response words can reflect prior processing (Ste-Marie & Jacoby, 1993), reduced distraction by suppressed (compared to baseline) response words can correspondingly index prior suppression. And because we used ruminative materials, we predicted a smaller reduction (less SIF) for participants who report habitual rumination.

## Method

**Materials.** Twenty-four cue-response pairs served as critical pairs to appear in the learning and TNT phases; the response words subsequently appeared as flankers in the indirect test. Although the cues were benign if considered individually, they were combined with emotionally negative response words to comprise pairs with ruminative potential (e.g., *body-insecure*, *party-loser*, *teacher-disappoint*, *dinner-quarrel*). Given the lack of relevant norms, the pairs were constructed merely from our own imagination, but with the constraints of available norms for the individual words. Three sets of 8 pairs were balanced on emotional-valence ratings from the Affective Norms for English Words (ANEW; Bradley & Lang, 1999), concreteness ratings (Brysbaert, Warriner, & Kuperman, 2014), and frequency (Kucera & Francis, 1967) of the cue and the response word. Sets were counterbalanced with the role assigned in the first two phases of the session: baseline (learned but not cued in the TNT phase), suppression (learned and then suppressed in the TNT phase), and new (neither learned nor cued in the TNT phase). Again, all of these critical pairs consisted of benign cues and negative response words. Sixteen additional benign cues with benign response words (e.g., *butter-pancakes*, *baseball-youth*) served as “think” items during the TNT phase. They were initially learned and practiced during the TNT phase (as the “think” items), but did not appear in the flanker task.

Figure 1 represents all types of trials in the flanker task; each flanker was paired with a new target, matched on word length. The flanker task required 56 targets to be judged for their valence and 32 new flanker words (in addition to the 24 critical, negative response words from the previous phases). For the critical and valence trials (the top and middle rows of Figure 1), neutral targets were chosen from the mid-range of ANEW and assigned to sets so that their mean valence ratings were also balanced across the three critical and two new sets (5.2 – 5.3 on the 9-

pt scale). Other details of flanker materials correspond exactly to those reported by Hertel and Hayes (2015). (Materials for both experiments are available on the Open Science Framework at DOI 10.17605/OSF.IO/N7ZV3.)

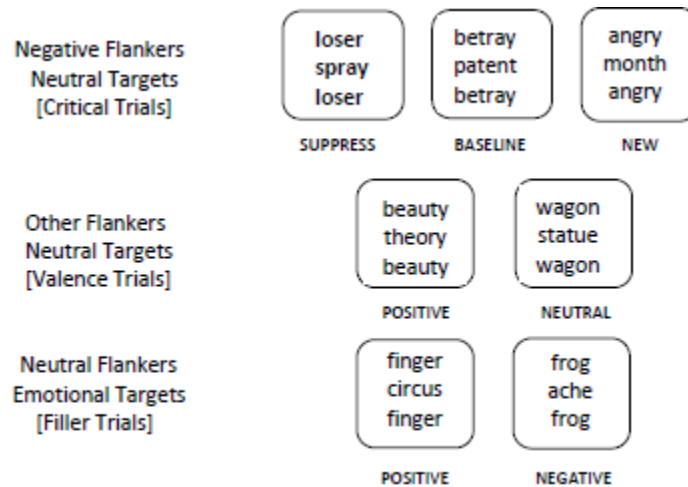


Figure 1. Examples of materials for the flanker task.

**Participants and design.** We planned to recruit 30 participants in each rumination group.<sup>2</sup> Multiples of 6 were required for counterbalancing the critical sets with the within-subjects factor of TNT role (suppress, baseline, and new) and, crossed with that factor, the direction of the rating scale during the flanker task (negative to positive, or vice versa).

To obtain those 60 participants, we screened introductory psychology classes at Trinity University with the RRS form and recruited students who scored in the top and bottom quartiles of the total score. By the end of recruiting, 83 students seeking course-required credit participated in what we described as two experiments—one designed for learning word pairs and another for assessing the emotional value of words. Twelve students' data were set aside because their end-of-session RRS scores placed them out of their original quartile by at least 5 points

<sup>2</sup> Previous TNT experiments in our lab and elsewhere have recruited 18-24 participants per condition. We increased the sample size due to the unknown characteristics of the grouping factor of rumination and the lack of a similar previous experiment to provide the details for a power analysis.)



(five from the first and seven from the fourth quartile). An additional 11 students' data were excluded due to noncompliance with suppression instructions in the TNT phase (see the procedure section). When identified, these participants were replaced so that, as intended, the final sample included 30 participants (19 women) in each rumination group. Counterbalancing was disturbed in both groups by one (and the same) assignment error. In the final sample, age ranged from 18 to 22. Forty-two percent identified at least in part as Latinx/Hispanic, 67% as Anglo/Caucasian, 12% as East or Southeast Asian, and 5% as African-American/Black. Mean BDI-II and RRS scores, collected at the end of the session, are reported in Table 1.

**Procedure.** SuperLab Pro (Version 5.0; Cedrus Corporation, San Pedro, CA) was used to program all phases. A 200-ms orientation display (two crosses separated by 10 spaces horizontally) initiated all trials in each phase; trials ended with a 500-ms ISI. All words or pairs were centered on the screen and displayed in Tahoma 20-pt regular font on white background; unless otherwise reported, word color was black.

**Learning phase.** Pairs from two critical sets (scheduled for baseline and suppression in the subsequent TNT phase), together with the 16 pairs scheduled for responding, were displayed for 5 s each. Participants were asked to create a mental image for the pair to help them learn it and to type a number to rate the vividness of the image on a 7-pt rating scale ranging from “not at all vivid” to “extremely vivid.” Pairs were randomized anew for each participant within blocks of eight, each block containing four pairs scheduled for responding in the next phase, two for suppression, and two for non-presentation (baseline). After presenting all 32 pairs, we tested learning by presenting the cues in the same randomized-block ordering and repeating the blocks in turn by including a cue only when the previous response to it had been incorrect. On each trial

of this test, the cue was presented for 5 s or until a response was spoken aloud; the cue was followed by a display of the correct response in blue for 2 s.

***TNT phase.*** In a brief practice task and the main TNT phase, a green cue word signaled the participant to practice recalling the associated response; corrective feedback was provided on error trials only (by the experimenter-initiated appearance of the correct response in blue). A red cue signaled participants to not think about the corresponding response word for the duration of the 3-s display, while nevertheless focusing attention on the cue. Direct-suppression instructions repeatedly emphasized the importance of preventing the response word or any other thought from coming to mind. Any response to a red cue was followed by a string of large red Xs, initiated by the experimenter's keypress. Trial order in the TNT phase was randomized within four blocks of six cues: four in green for responding and two in red for suppression. All blocks were presented 16 times in succession, with a break after 8. (Thus, each cue was presented 16 times.) The task ended with a questionnaire to assess compliance with suppression instructions (see Hertel & Calcaterra, 2005). From the participants' point of view in our cover story, however, the questionnaire served as our main method of assessing their ability to not think about words they had learned. They understood that the experimenter would not see their responses.

***Flanker task.*** In the next "experiment," according to the cover story, participants were asked to rate the emotional valence of the middle word of a 3-word, vertically aligned display; they were told that the word presented above and below it was unimportant; it would be rated by other participants. Ratings were made "as quickly as possible" (at any point during the display) on a scale from 1 to 5, where 1 meant either very negative or very positive (according to the counterbalancing scheme) and 5 meant the opposite. Using a *QWERTY* keyboard, participants

positioned the first two fingers of each hand on the keys *C* (1 or 5), *V* (2 or 4), *N* (4 or 2), and *M* (5 or 1); the rating of 3 was made by using either index finger to press *B*. Following a practice block of 6 buffer pairs (all new words) completed in the presence of the experimenter, the 56 flanker pairs were presented. Their order was randomized within eight blocks of 7 events, one from each type illustrated in Figure 1.

At the end of the session, participants completed the RRS, the Beck Depression Inventory (BDI-II; Beck, Steer, & Brown, 1996), and a short form for indicating their age and race or ethnicity. All paper forms were placed in an envelope and sealed by the participant, who understood that the experimenter would not have access to these data.

## Results

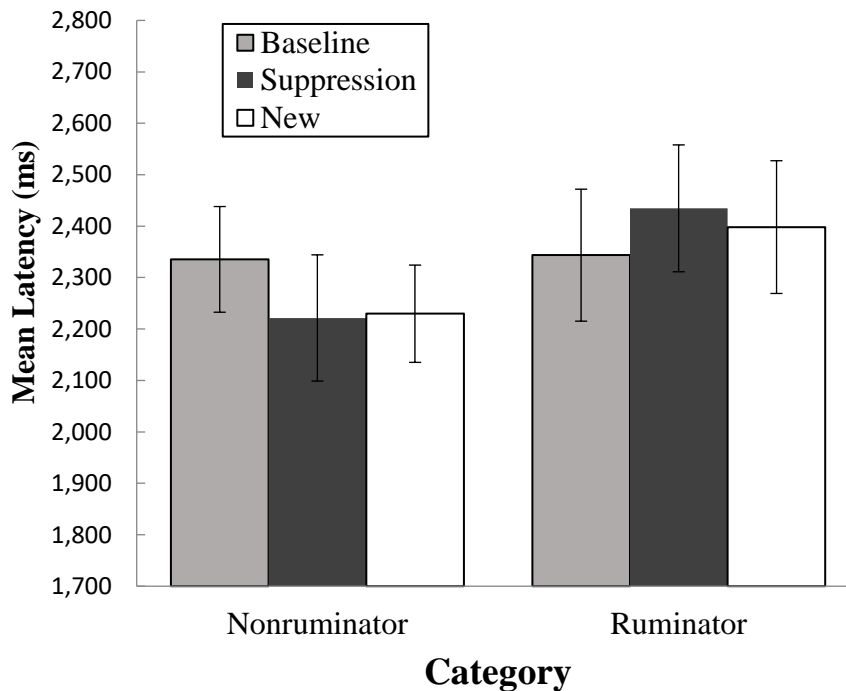
**Flanker effects.** The primary dependent measure was the time in ms to rate the valence of the target word, recorded as the mean latency across the eight observations of each type. Mean latencies were submitted to a mixed-design analysis of variance (ANOVA) with a between-subjects factor for rumination category (nonruminator or ruminator) and a within-subjects factor for flanker status in the critical trials (baseline, suppressed, and new). Means are illustrated in Figure 2.<sup>3</sup>

The interaction between rumination category and flanker status was marginally significant,  $F(2, 116) = 2.90$ ,  $MSE = 60088$ ,  $p = .059$ ,  $\eta_p^2 = .05$ . Planned interaction comparisons involved the baseline flankers compared separately to the new and suppressed flankers. First, the prior-processing effect (baseline vs. new) did not significantly depend upon rumination category,

---

<sup>3</sup> Other flanker trials—beyond the critical trials—were included to mask our interests and provide emotional words as targets to rate. The design of trial type permitted the analysis of whether flanker valence (new negative, positive, and neutral) was differentially distracting. It was not (see also Hertel & Hayes, 2015).

$F(1, 58) = 3.06$ ,  $MSE = 62882$ ,  $p = .085$ ,  $\eta_p^2 = .05$ . And the effect itself was nonsignificant even for nonruminators,  $t(29) = 1.31$ ,  $p = .201$ , 95%  $CI_{diff} = [-59.5, 270.8]$ . More central to our concern was the significant interaction of the suppression effect (baseline vs. suppression) with rumination category,  $F(1, 58) = 5.80$ ,  $MSE = 54222$ ,  $p = .019$ ,  $\eta_p^2 = .09$ . As shown in Figure 2, target ratings by nonruminators were slowed less by suppressed flankers compared to baseline flankers,  $t(29) = 2.42$ ,  $p = .022$ , 95%  $CI_{diff} = [17.7, 210.1]$ . Target ratings by ruminators were not,  $t(29) = -1.28$ ,  $p = .209$ , 95%  $CI_{diff} = [-235.8, 53.9]$ .



**Figure 2.** Mean latency in ms to rate the emotional valence of the target when flanked by emotionally negative response words from the previous TNT phase: learned but not practiced responses (baseline), learned and suppressed responses (suppression), or words encountered for the first time (new). Error bars represent standard errors.

**Valence ratings.** Because the 5-pt rating scale is less sensitive to differences than the scale used in Experiment 1 by Hertel and Hayes (2015), we did not expect the prior status or the valence of the flankers to affect the ratings of the central (neutral) targets. However, we did

analyze the ratings of the positive, neutral, and negative targets, flanked by neutral words, as a measure of whether participants stayed on task. Clearly, they did, as indicated by the main effect of target valence,  $F(2, 116) = 358.23$ ,  $MSE = .353$ ,  $p < .001$ ,  $\eta_p^2 = .86$ . Means were 1.5 (negative targets), 3.0 (neutral), and 4.5 (positive). The interaction with rumination group was nonsignificant,  $p = .67$  and  $\eta_p^2 = .01$ .

**Self-report measures.** Noncompliance with suppression instructions was measured by the mean rating for three items on the questionnaire following the TNT phase. A rating of 0 indicated that the participants never used a particular strategy to remember the suppressed words; 1 indicated rarely, 2 sometimes, 3 frequently, and 4 very frequently. The 11 participants whose data were replaced produced means greater than 1 (our cut-off in this and the two prior experiments by Hertel & Hayes, 2015). Noncompliance did not depend on rumination group for the remaining participants,  $M = 0.40$  in each group.

Finally, average BDI-II and RRS scores (end of session) are reported in Table 1. The RRS confirmed the initial selection difference. The two groups also differed on the BDI-II,  $t(42.32) = 6.15$ ,  $p < .001$ , 95%  $CI_{diff} = [8.5, 16.8]$ .<sup>4</sup>

## Experiment 2

The results from Experiment 1 provided indirect evidence for the forgetting of suppressed response words in the TNT paradigm. Nonruminators were distracted less by previously suppressed flankers than baseline flankers, whereas ruminators responded similarly to the two types of trials. Although the results clearly show that the SIF effect was absent for ruminators, their similar latencies across the three types of flankers (baseline, suppression, new) suggest that

---

<sup>4</sup> We administered a form of the BDI-II that omitted the suicide item; the score is therefore summed across 20 items, not 21. One participant in each group failed to complete the form.

negatively valenced words more generally distracted them during the rating task. We therefore turned our attention to a different paradigm for obtaining indirect evidence of forgetting.

Hertel, Large, Stück, and Levy (2012) developed a free-association task to serve as an indirect measure of forgetting in the TNT paradigm. Cues used during the learning and TNT phases were intermixed with other cues in a purportedly separate task to collect free-association data. Although there was no guarantee that participants did not catch on and treat the task as a cued-recall task, end-of-session ratings for the deliberate production of words from the TNT experiment were not correlated with the main measure of forgetting. Even in the few instances of reported awareness of the produced word's prior role, it is difficult to know whether awareness had been experienced before or after the word was produced (deliberately or coincidentally). Therefore, the free-association task seemed viable for our purpose.

Using free association as an indirect test is not as simple as it might seem at first. If the free-association response is often a response word from a prior experimental phase, participants are likely to catch on and turn the indirect test of memory into a direct test. To counteract that possibility, we developed word pairs consisting of a homographic cue and a meaningfully related response word that is nevertheless a normatively low associate of the cue (e.g., *glare-grudge*, *image-ruined*). Lacking sufficient numbers of rumination-related homographs, we also developed nonhomographic cues that could take on negative or benign meanings, depending on the response word (e.g., *choice-mistake*, *perform-choke*). By designing these materials, we made it possible to score the produced responses in the free-association task according to whether they reflected the meaning of the cue that was established by the response word during the TNT phases, or an alternative, more benign meaning that is also generated more frequently (Nelson, McEvoy, & Schreiber, 1998). We call this method the scoring of learned-meaning responses.

Other than these considerations concerning the indirect test, the design and method of Experiment 2 was very similar to Experiment 1.

## Method

**Participants.** Students in introductory psychology and psychology-statistics courses at Trinity University participated for course credit or 10 dollars compensation. Students were recruited according to their rumination category (see Experiment 1), and within the two categories they were randomly assigned to one of three counterbalancing conditions for rotating sets of word pairs across conditions of the within-subjects factor for TNT status (baseline, suppression, and new). We planned a sample size of 27 in each rumination category (vs. 24 in the experiment by Hertel et al., 2012). To that end, we recruited 74 participants; the data from 7 students were set aside due to their noncompliance scores on the TNT questionnaire, 11 because their RRS scores were unstable (similarly to Experiment 1), 1 due to the failure to provide RRS scores at session end, and 1 assignment error. In the final sample, 4 of the 9 participants in each counterbalancing cell were male, and 5 were female. Twenty-eight percent identified at least in part as Latinx/Hispanic, 57% as Anglo/Caucasian, 17% as East or Southeast Asian, 4% as another race/ethnicity, including African-American. Ages ranged from 18 to 24. Mean BDI-II and RRS scores collected at the end of the session are reported in Table 1.

**Materials.** We constructed 18 cue-response pairs, 9 of which contained homographs as cues paired with negative response words (e.g., *blue-gloom*) and 9 of which contain cues with potentially negative meanings for ruminators—meanings that we invoked with our choice of response words (e.g., *essay-fail*). We distributed the pairs into three sets of 6 pairs (3 homographic) and balanced the sets most strictly on the rated valence of the pairs in pilot work, the forward associative strength (fsg, Nelson, et al., 1998) to the most frequent (and always

benign) associate, and the summed fsg to all associates with benign meanings and all associates with negative meanings. The sets were also balanced, more loosely, on the frequency, valence, and concreteness of the separate cues and targets. Two of the sets appeared during the learning and TNT phases, along with 12 benign pairs (chosen from the ones used in Experiment 1). Thus, only 24 pairs were studied in the learning phase, and only 18 cues were encountered (repeatedly) during the TNT phase. This reduction was accomplished in order to minimize the likelihood of a recollective approach to the free-association task. In that task, participants encountered just 12 old cues, along with 6 new cues from the remaining critical set and 30 new filler cues, chosen to resemble the critical cues in valence, concreteness, frequency, and number of letters. Thus, one-fourth of the cues were old.

**Procedure.** Apart from the reduced number of critical items, the learning and TNT phases were the same as in Experiment 1. And, as in Experiment 1, the TNT phase ended with the questionnaire to assess compliance with suppression instructions and to provide a mock ending to the experiment. After their responses were sealed, the experimenter delivered a cover story: “The purpose of this [next task] is to get normative data for words that will appear in an experiment next fall. There might be a little overlap in words—I don’t know—but this task is entirely different.” The last sentence was delivered nonchalantly and followed by informal chat about how we added this task to the experiment they finished to avoid the struggle to find participants for such a brief study. Then the experimenter instructed the student to speak aloud the first word that comes to mind in response to each word presented on the monitor; more than one word could be reported if they seemed to occur simultaneously.

The test began with two practice blocks containing 6 filler cues each and, after a brief pause to emphasize instructions, continued with six blocks of 6 cues each: 2 old cues (1 that cued



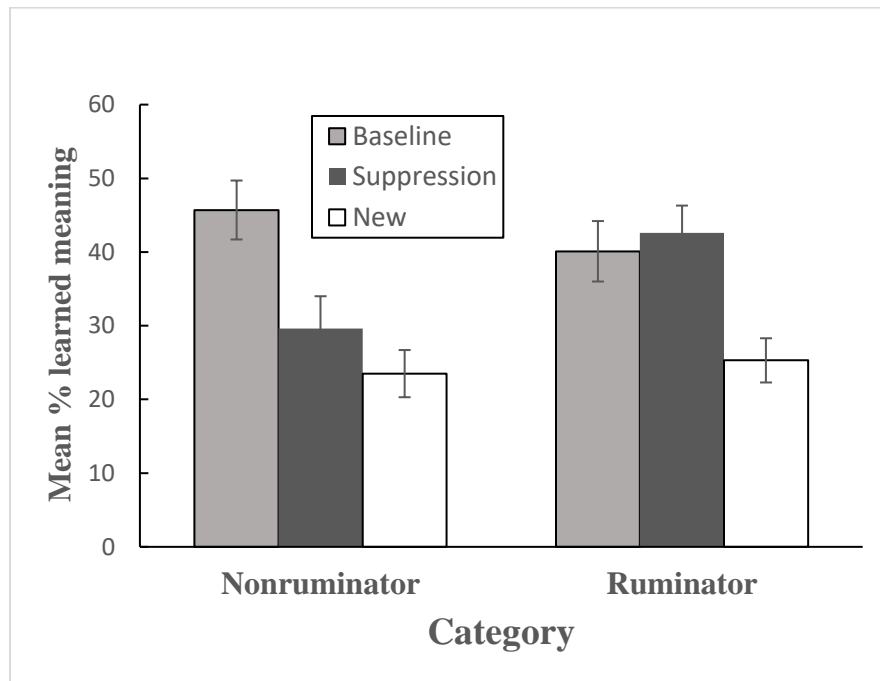
a suppressed response in the TNT phase and 1 baseline) and 4 new cues (1 from the remaining critical set and 3 fillers). Assignment of critical cues to blocks mimicked the assignment in the earlier phases, and order within blocks was randomized anew for each participant. Each cue was presented for 3.5 s. If participants responded too slowly, a “Speed Up” sign appeared in the center of the screen for 600 ms to encourage faster responding. Participants responded aloud.

At test completion, we administered another questionnaire, in which participants were asked to reflect about their reactions to the free-association task (“How often did you...”). It consisted of seven items, only two of which were relevant to our purpose: Item 5 (“...deliberately respond with words from the experiment”) and Item 7 (“...respond quickly off the top of your head”). The other items were included to mask our interest. Responses were made on a scale of 0-4, from *never* to *very frequently*. Then, participants filled out the RRS, the BDI-II, and a demographic form.

## Results

**Free-association responses.** Two judges, uninformed about the ruminative condition of the participants, independently categorized the 18 responses to cues from the three critical sets (baseline, suppressed, and new) according to whether they connoted the learned negative meaning of the cue. They agreed about 94% of the responses, and the disagreements were resolved by a third judge. The percentages of learned meanings were then submitted to a mixed-design analysis of variance, with a between-subjects factor for rumination category (nonruminators or ruminators) and within-subjects factors for cue status (baseline, suppressed, and new) and cue type (homograph or not). To reduce error variance, a second between-subjects factor was included to represent the factor for rotating sets across the levels of cue status; significant effects associated with this counterbalancing factor are not reported. Nor are

significant lower-order effects that are qualified by higher-order effects. Homographic cues produced more responses connoting learned meanings ( $M = 42.2$  vs.  $26.7$  for nonhomograph cues),  $F(1, 48) = 24.86$ ,  $MSE = 775.92$ ,  $p < .001$ ,  $\eta_p^2 = .34$ . Because this factor did not interact with other factors in the design (all  $p$  values  $> .25$ ), we combined the data from the two types in subsequent analyses. Figure 3 illustrates these combined means as a function of cue status and rumination category.



*Figure 3.* Mean percentages of responses indicating the learned meanings of the cues (both types combined), depending on the cue's role in the earlier TNT phases: its response learned but not practiced subsequently (baseline), its response learned and then repeatedly suppressed, or the cue was never encountered (new). Standard error bars depicted.

As shown in Figure 3, the effect of prior experience with the cues depended on the participants' rumination category,  $F(2, 96) = 3.78$ ,  $MSE = 310.40$ ,  $p = .026$ ,  $\eta_p^2 = .07$ . The interaction of category with the comparison between baseline and new cues was nonsignificant,  $p = .246$ . Instead, the overall interaction is better explained by the interaction comparison between baseline and suppressed conditions,  $F(1, 48) = 6.20$ ,  $MSE = 746.04$ ,  $p = .016$ ,  $\eta_p^2 = .11$ . Clear

evidence of SIF was obtained for nonruminators,  $F(1, 24) = 10.94$ ,  $MSE = 635.54$ ,  $p = .003$ ,  $\eta_p^2 = .31$ ; but not for ruminators,  $p > .66$ . Ruminators more often interpreted the cues for suppression as negative, similarly to how all participants interpreted the cues from the baseline condition.

**Questionnaire reports of strategy use.** The suppression effect was not significantly correlated with responses to Items 5 and 7 on the final questionnaire. The direction of the correlation for Item 7 (respond quickly) suggested that participants who believed they responded quickly were apt to show more forgetting,  $r(52) = .22$ ,  $p = .118$ . If they thought they were deliberately remembering (Item 5), the direction of the correlation suggested smaller suppression effects,  $r(52) = -.17$ ,  $p = .214$ . Again, however, the correlations were nonsignificant. Moreover, there were nonsignificant differences between ruminators and nonruminators in responses to these items, each  $p > .57$ .

### General Discussion

The phenomenon of rumination is an interesting problem for applied experimental psychology, because ruminators constitute a group of people with an identifiable cognitive habit. Cognitive-behavioral therapists quite obviously focus on cognitive habits, and they frequently observe the difficulties involved in stopping thoughts. Our results confirm difficulties associated with the habit of rumination. In two experiments, we arranged for repeated suppression practice in the presence of cues for the associated memories. Nonruminators' responses in the subsequent transfer tasks revealed that suppression practice can be beneficial, at least in the short term, but ruminators failed to show those benefits. In similar work, Fawcett et al. (2015) had previously documented rumination-related deficits in SIF on tests of deliberate recall of neutral materials. Our results were obtained from more ecologically relevant tests and with materials related to the

tendency to ruminate. We next discuss possible ways to understand why rumination-related deficits occur.

First, the ruminative deficit seems rather obviously to be connected to poor attentional control. Although the participants were selected according to scores on a rumination inventory, as is typical, they were also dysphoric and possibly depressed.<sup>5</sup> And the prevalent explanation for cognitive difficulties associated with depressive rumination is based on demonstrations of poor attentional control (see Koster, De Lissnyder, Derakshan, & De Raedt, 2011; e.g., Quinn & Joormann, 2015). Regardless of whether the performance deficit occurs in measures of attending, remembering, or forgetting, inferences about control rely on the extent to which task performance requires unassisted focus (i.e., the extent to which performance requires self-control; Hertel, 2004). Many of these tasks—in particular, the ones utilizing ruminatively relevant materials—also invite questions about the performance/competence distinction. Do the depression/rumination-related differences result from fundamental deficits in the ability to control attention (the prevalent claim), or do they reflect variations in interest, motivation, or some combination? Ruminators are interested in events (perhaps even experimental events) related to their own personal concerns. Therefore, when attention is not well controlled experimentally or environmentally, ruminators practice thinking about those concerns and develop habits that, in new situations, compete with the behaviors that presumably imply good cognitive control. From this point of view, then, what is often interpreted as a fundamental

---

<sup>5</sup> Scores on the BDI-II cannot be used meaningfully as covariates in either experiment, given their strong correlation with the RRS scores comprising the grouping variable (see Miller & Chapman, 2001). It is possible to ignore RRS scores and divide the scores on the dependent variables into high and low BDI scores within each counterbalancing cell and then rerun all analyses. However, in this case, given the BDI / RRS correlations and our use of first and fourth quartiles during recruiting, the groups would be comprised of almost if not exactly the same participants as in the reported results. Such divisions, moreover, can mean that someone with a BDI-II score of 11 (for example) is placed in the low group and someone with a 12 in the high group. This procedure would not produce a good basis for concluding anything about dysphoria and forgetting; “middle scores” are usually omitted in clinical cognition research on depression or dysphoria.

deficit might simply be a well-developed habit, established and maintained by interest. It also seems worth mentioning that ruminators might be more motivated to remember than to suppress or forget, in the interest of resolving problems and concerns (Nolen-Hoeksema, Wisco, & Lyubomirsky, 2008). Alternatively, there is some evidence that individuals exhibit control deficits prior to characterizing themselves as ruminators (Koster et al., 2011). Although the causal connection is difficult to establish, these deficits have been shown to contribute to ruminative thinking (see Hoorelbeke, & Koster, 2017; Mor & Daches, 2015, for a review) and may contribute to the onset of ruminative habits. In either case, it now seems unwise to expect “direct” suppression from a ruminator, as any competent therapist already knows.

Instead of demanding self-control, practicing competing responses to cues for rumination might be efficacious in changing ruminative habits—by replacing an old habitual thought with a new one. To break a habit in everyday situations or in therapeutic contexts, people frequently and repeatedly invoke competing responses. In the laboratory, experimental procedures of thought substitution have facilitated SIF for both dysphoric and non-dysphoric students (using neutral materials, Hertel & Calcaterra, 2005) and depressed and nondepressed adults from the community (using negative materials, Joormann, Hertel, LeMoult, & Gotlib, 2009).<sup>6</sup> More recently, a retrieval-practice experiment performed with rumination-related word pairs showed that ruminators recalled the negative targets that were not even practiced when the test was immediate, and they recalled them as well as everyone recalled practiced words; a week later, however, that bias had disappeared and even the ruminators’ recall revealed only the benefit of having practiced retrieval of emotionally positive pairs (Hertel, Maydon, Cottle, & Vrijzen, 2017). All of this evidence about the benefits of positive substitution, however, was collected on

---

<sup>6</sup> Substitution procedures likely operate by establishing retrieval-induced forgetting (Hertel & Calcaterra, 2005).

direct tests of memory. Substitution during suppression practice does transfer to reveal strong SIF during free-association (with neutral materials; Hertel et al., 2012). Practicing positive thoughts that compete with ruminative concerns is therefore a potential method for indirectly suppressing negative thoughts to aid their forgetting.

In the meantime, however, we stress the importance of varying the conditions for suppression or substitution practice (Bouton, 2000) and the importance of indirect tests. Direct tests likely cue the practice situation (TNT in this case) and thereby provide unrealistically hopeful data for suppression-induced forgetting in the real world. Indirect tests simulate common conditions in which negative thoughts can be inadvertently cued, and they are thereby likely to be more sensitive to generalization from original learning.

### References

- Anderson, M. C., & Huddleston, E. (2012). Towards a cognitive and neurobiological model of motivated forgetting. In R. F. Belli (Ed.), *True and false recovered memories: Toward a reconciliation of the debate* (Nebraska Symposium on Motivation, Vol. 58, pp. 53–120). New York: Springer.
- Beck, A. T., Steer R. A., & Brown, G. K. (1996). Manual for the Beck Depression Inventory (2<sup>nd</sup> Ed.). San Antonio, TX: Psychological Corporation.
- Bjork, R. A. (1989). Retrieval inhibition as an adaptive mechanism in human memory. In H. L. Roediger and F. I. M. Craik (Eds.), *Varieties of memory and consciousness: Essays in honour of Endel Tulving* (pp. 309-330). Hillsdale, NJ: Erlbaum.
- Bouton, M. (2000). A learning theory perspective on lapse, relapse, and the maintenance of behavior change. *Health Psychology, 19*, 57-63.
- Bradley, M. M., & Lang, P. J. (1999). *Affective norms for English words (ANEW): Instruction manual and affective ratings*. Technical Report C-1, The Center for Research in Psychophysiology, University of Florida.
- Brysbaert, M, Warriner, A. B., & Kuperman, V. (2014). Concreteness ratings for 40 thousand generally known English word lemmas. *Behav Res 46*:904–911. DOI 10.3758/s13428-013-0403-5
- Cooney, R. E., Joormann, J. Eugene, F., Dennis, E. L., & Gotlib, I. H. (2010). Neural correlates of rumination in depression. *Cognitive, Affective, & Behavioral Neuroscience, 4*, 470-478.
- Fawcett, J. M., Benoit, R. G., Gagnepain, P., Salman, A., Bartholdy, S., Bradley, C., & Anderson, M. C. (2015). The origins of repetitive thought in rumination: Separating cognitive style

from deficits in inhibitory control over memory. *Journal of Behavior Therapy and Experimental Psychiatry*, 47, 1-8.

Hertel, P. T. (1998). Relation between rumination and impaired memory in dysphoric moods. *Journal of Abnormal Psychology*, 107, 166-172.

Hertel, P. T. (2004). Memory for emotional and nonemotional events in depression: A question of habit? In D. Reisberg & P. Hertel (Eds.), *Memory and emotion* (pp. 186-216). New York: Oxford University Press.

Hertel, P. T., & Calcaterra, G. (2005). Intentional forgetting benefits from thought substitution. *Psychonomic Bulletin & Review*, 12, 484-489.

Hertel, P.T., & Gerstle, M. (2003). Depressive deficits in forgetting. *Psychological Science*, 14, 573-578.

Hertel, P. T., & Hayes, J. A. (2015). Distracted by cues for suppressed memories. *Psychological Science*, 26, 775-783.

Hertel, P. T., Large, D., Stück, E. D., & Levy, A. (2012). Suppression-induced forgetting on a free-association test. *Memory*, 20, 100-109.

Hertel, P., Maydon, A., Cottle, J., & Vrijssen, J. (2017). Cognitive bias modification: Retrieval practice to simulate and oppose ruminative memory biases. *Clinical Psychological Science*, 5, 122-130.

Hoorelbeke, K., & Koster, E. H. (2017). Internet-delivered cognitive control training as a preventive intervention for remitted depressed patients: Evidence from a double-blind randomized controlled trial study. *Journal of Consulting and Clinical Psychology*, 85, 135.



- Hue, X., Bergstrom, Z. M., Gagnepain, P., & Anderson, M. C. (2017). Suppressing unwanted memories reduces their unintended influences. *Current Directions in Psychological Science*, *26*, 197-206.
- Joormann, J., Hertel, P. T., LeMoult, J., & Gotlib, I. H. (2009). Training forgetting of negative material in depression. *Journal of Abnormal Psychology*, *118*, 34-43.
- Joormann, J., Levens, S. M., & Gotlib, I. H. (2009). Sticky thoughts: Depression and rumination are associated with difficulties manipulating emotional material in working memory. *Psychological Science*, *22*, 979-983.
- Kim, K., & Yi, D.-J. (2013). Out of mind, out of sight: perceptual consequences of memory suppression. *Psychological Science*, *24*, 569-574.
- Koster, E.H.W., De Lissnyder, E., Derakshan, N. & De Raedt, R. (2011). Understanding depressive rumination from a cognitive science perspective: The impaired disengagement hypothesis. *Clinical Psychology Review*. *31*, 138-145.
- Kucera, H., & Francis, W. N. (1967). Computational analysis of present-day American English. Providence, RI: Brown University Press.
- Miller, G. A., & Chapman, J. P. (2001). Misunderstanding analysis of covariance. *Journal of Abnormal Psychology*, *110*, 40-48.
- Miller, R. R., & Laborda, M. A. (2011) Preventing recovery from extinction and relapse: A product of current retrieval cues and memory strengths. *Current Directions in Psychological Science*, *20*, 325-329.
- Mor, N., & Daches, S. (2015). Ruminative thinking: lessons learned from cognitive training. *Clinical Psychological Science*, *3*, 574-592.

Nolen-Hoeksema, S., Wisco, B. E., & Lyubomirsky, S. (2008). Rethinking rumination.

*Perspectives on Psychological Science, 3*, 400–424.

Quinn, M. E., & Joormann, J. (2015). Stress-induced changes in executive control are associated with depression symptoms: Examining the role of rumination. *Clinical Psychological Science, 3*, 628-636.

Ste-Marie, D. M., & Jacoby, L. L. (1993). Spontaneous versus directed recognition: The relativity of automaticity. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19*, 777-788.

Treynor, W., Gonzalez, R., & Nolen-Hoeksema, S. (2003). Rumination reconsidered: A psychometric analysis. *Cognitive Therapy and Research, 27*, 247-259.

Watkins, E. R., & Nolen-Hoeksema, S. (2014). A habit-goal framework of depressive rumination. *JAbn, 123*,24-34.

Yang, Y., Cao, S., Shields, G. S., Teng, Z., & Liu, Y. (2017). The relationships between rumination and core executive functions: A meta-analysis. *Depression and Anxiety, 34*, 37-50.

Table 1

*Means (Standard Deviations) for Self-Report Measures in Both Experiments*

	Experiment 1		Experiment 2	
	Nonruminators	Ruminators	Nonruminators	Ruminators
RRS	35.6 (5.84)	62.6 (7.47)	34.5 (6.28)	62.7 (7.59)
BDI-II	7.6 (5.13)	20.2 (9.78)	6.8 (4.11)	19.3 (9.56)

*Note.* RRS = Ruminative Response Scale, BDI-II = Beck Depression Inventory.

