10-2021

Rumination: Practicing Retrieval of Autobiographical Memories

Paula T. Hertel
Trinity University, phertel@trinity.edu

Paola Herrera
Trinity University, pherrer1@trinity.edu

Pallavi Shamapant
Trinity University, pshamapa@trinity.edu

Publication Details
Cognitive Therapy and Research

Rumination: Practicing Retrieval of Autobiographical Memories

The final publication is available at https://link.springer.com/article/10.1007/s10608-020-10118-z

Trinity University

Author Note

The authors are grateful to many students in the lab: Amaris Maydon for help with materials, Sarah Gorban, Victoria Marshall, Sierra Rodriguez, and Mariale Torres for scoring, Etienne Ramirez-Gonzales for conducting some sessions, and Samantha Wilkerson for the collection of pilot data.

Abstract

Background: People who ruminate about negative personal experiences seem to be engaged in practicing retrieval, with the expected consequences of perpetuating those very memories (see Roediger & Butler, 2011).

Method: To provide an experimental model of the effect of retrieval practice on subsequent recall of autobiographical memories, we recruited students with low and high scores on the Ruminative Response Scale (Treynor, Gonzalez, & Nolen-Hoeksema, 2003), gave them positive and ruminative cues to generate memories, and cued three rounds of practice of half of the memories from each valence.

Results: A week later, final cued recall for all memories showed strong practice effects, with the exception that ruminators benefited very little from the request to practice positive memories.

Conclusion: Recalling personal memories benefits from prior practice in bringing them to mind. However, these results provide insufficient evidence to recommend training recall of positive autobiographical memories to counteract rumination.

Rumination: Practicing Retrieval of Autobiographical Memories

Personal experience suggests that the more often we think a thought, the more often we continue to think that thought; the thought becomes a habit. This aphorism has been expressed in philosophical psychology of old (James, 1890), and it crudely represents the more modern science of transfer-appropriate processing (Morris, Bransford, & Franks, 1977) and retrieval practice (Bjork, 1975, Roediger & Butler, 2011), as well as the clinically relevant phenomena called rumination, worry, and repetitive thinking. Thus, the persistence of practiced thought has captured attention across subdisciplines of psychology.

Rumination is a term that refers to the largely self-focused, repetitive, and negative mental experiences that typify depressed and anxious individuals (Nolen-Hoeksema, Wisco, & Lyubomirsky, 2008; Watkins & Nolen-Hoeksema, 2014). An early classic study of rumination and autobiographical memory (Lyubomirsky, Caldwell, & Nolen-Hoeksema, 1998) revealed that scores on a standard rumination instrument were positively correlated with the production of negatively biased personal memories. Ruminators remembered negative events, possibly because they had practiced retrieving them. During the years since that ovular contribution, a wide stream of research has dedicated itself to discovering the cognitive characteristics of rumination, sometimes with the goal of reducing the tendency to ruminate in the service of treating anxious or depressed individuals. Deficient cognitive control has emerged as a clear hallmark of ruminative tendencies (see the meta-analysis by Zetsche, Burkner, & Schulze, 2018); for example, ruminators do not easily switch attention from negative experimental material (Joormann & Gotlib, 2008), but some evidence suggests that they can be trained to do so in ways that reduce their ruminative tendencies (Cohen, Mor, & Henik, 2015). Similarly, compared to others, ruminators show smaller effects of suppression practice and directions to forget on

subsequent memory tasks (Fawcett et al., 2015; Hertel, Maydon, Ogilvie, & Mor, 2018; Joormann & Tran, 2009). They are also quick to interpret ambiguous events negatively. That tendency contributes to rumination, given experimental evidence that training a ruminative interpretation bias creates negatively biased recall of new ambiguous situations *and* higher rumination scores (Hertel, Mor, Ferrari, Hunt, & Agarwal, 2014). The good news is that training ruminators to ignore negative material can reduce subsequent memory bias (Daches, Mor, & Hertel, 2019). These findings on deficient control, habitual biases, and efforts to correct both are just a sampling of the wide literature on rumination and cognition, but evidence thus far had not connected rumination directly to the practice of repeated retrieval.

The aspect of ruminative remembering that interests us is repetitive focus on specific past events, not merely on a larger number of negative events (as in mood-congruent memory), but the same event(s) retrieved repeatedly. Recently, Hertel, Maydon, Cottle, and Vrijsen (2017) brought the science of retrieval practice to bear on the question of that persistence (also see Vrijsen, Hertel, & Becker, 2016). The experiment consisted of a learning phase in which ruminative and other students studied word pairs with negative, neutral (as fillers), and positive connotations. The negative pairs represented concerns that these students might have developed (e.g., *embarrassing body*), whereas the positive pairs represented happier experiences (e.g., *praised paper*). In each of three rounds in the next phase, the participants studied all materials and then practiced recalling the nouns when cued with the adjectives from the neutral pairs and either the negative or the positive pairs (or they studied without retrieval practice). At the end of the session, an immediate test cued recall of half of the nouns of each emotional valence (practiced and not), and then all nouns were tested two days later. The interesting combination of outcomes in this experiment was that non-practiced nouns showed a ruminative bias as strong as

any practice effect on the immediate test, but two days later only the practice effects remained (and they were substantial). The authors suggested that practicing positive memories might thereby provide a mechanism for counteracting the habitual bias that so clearly perpetuates rumination and contributes to depression.

The obvious next step was to conduct a similar study with autobiographical memories instead of word pairs. First, however, it was important to discover whether practice effects could obtain in the realm of autobiographical memory and, secondarily, whether the practice effects would be emotionally congruent or incongruent with the participants' naturally occurring habits of thought. The tendency to think habitually about negative memories, for example, might suggest that corresponding practice effects would be smaller because even the unpracticed items might be memorable for other reasons. On the other hand, ruminators might be less inclined to practice positive memories along with negative memories, and this outcome would suggest difficulties in using positive memories to overcome ruminative tendencies. These are the issues we addressed in the current experiment.

Whenever autobiographical memories are sought in the laboratory, cues must first be established. This procedure is not very different from what happens naturally, as memories are brought to mind in the real world. Someone might mention a paper that she wrote recently and then the self-focused listener might involuntarily recall a recent similar experience. We therefore provided cues in the generation phase of the experiment and again later during practice and on the tests. However, during generation (only), we manipulated the emotional valence of the generated memory by pairing the cue with an adjective (*failing paper* for some, *praised paper* for others). Again, we expected to find practice effects on a delayed test, and we were curious

about whether they would be qualified by the relation of rumination to the emotional nature of the original cues and corresponding memories.

Method

Overview

The phases of Session 1 consisted of memory generation, three rounds of retrieval practice for half of the memories, and an immediate cued-recall test for an orthogonal half of the memories. One week later we sent participants a link to a Qualtrics-implemented test of all memories initially generated. The procedure conceptually replicated those used by Hertel et al. (2017), with some important exceptions. First, instead of assigning participants to practice targets in response to either positively or negatively emotional cues, all participants practiced memories generated to both types of cues. Although this change meant that we could not examine effects of practicing just one type of memory, it was necessary, given the practical considerations of data collection at a small institution, to discover first whether practice effects would generalize to autobiographical memories. Another main change pertained to the nature of the cue. The experiment by Hertel et al. cued recall of the same sets of target words by all participants, regardless of the cue valence. For example, some participants learned to respond with *paper* to the cue *praised*, whereas others responded with *paper* to *failing*. In the current experiment, either *praised paper* or *failing paper* served as the cue to generate a memory, but subsequently only *paper* cued retrieval of the personal memory during retrieval practice and tests. These procedures represented our attempt to control characteristics of the recalled words or (in the current experiment) the type of event remembered, while experimentally varying emotional valence. The third feature of the current procedure that deserves attention is the dual aspect of recall. To keep the practice phase less time consuming, we asked participants to

practice thinking about the full memory but to merely report a title to identify it. Therefore, participants initially titled each of their personal memories and then recalled those titles after the full memories when tested. Finally, we tested after a week's delay, instead of a two-day delay.

Participants and Design

We recruited participants from the pool of students enrolled in Principles of Psychology across four semesters. In the absence of a prior similar method to use for determining sample size, we chose a cell size of 32 (as a factor of 8 for counterbalancing[1]). Hertel et al. (2017) had used a cell size of 20).

In class, students filled out copies of the Ruminative Response Scale of the Response Styles Questionnaire (RRS; Treynor, Gonzales, & Nolen-Hoeksema, 2003); those who scored in the first and fourth quartiles were invited to participate in categories we called nonruminators and ruminators, respectively. Within each group, we assigned four students to each of the eight counterbalancing conditions for determining which sets of three cueing nouns were assigned to each valence (e.g., *praised paper* vs. *failing paper*), each practice status (practiced or not), and each condition of the immediate test (tested or not tested immediately); three of the four students in each cell were female, and the other one male. Assignment was random before nearing the equal-n constraint. When data were set aside, another participant was chosen randomly from those who met criteria to be assigned to the same counterbalancing condition.

Data were set aside and replaced if students' RRS scores at the end of the first session placed them more than three points outside their original rumination group (6 from the nonruminators and 15 from the ruminators).[2] We also replaced data from students who did not

[1] There were actually two counterbalancing factors; see the Materials section.

[2] Our experience suggests that undergraduate students do not show very stable scores on the RRS. Our practice in all experiments concerning rumination has always been to set aside data from individuals who do not produce a similar

take the Qualtrics test during the seventh or eighth day following Session 1 (5 nonruminators and 3 ruminators), as well as data from students who misunderstood their instructions for the delayed test (5 nonruminators and 2 ruminators). Some loss from incomplete data is inevitable when a test is delivered via Qualtrics, a week after in-lab participation, even with a subsequent reminder. The misunderstood instructions grew out of our apparently successful cover story, concocted to discourage more retrieval practice during the week before the delayed test. We told them that we would be asking for new memories on the Qualtric form, and failing to read test instructions these participants produced only new memories. Finally, incomplete data were obtained from an additional six students, due to freezes in the program or experimenters choosing the wrong program. We report data from 64 students, 32 in each rumination group (24 women and 8 men per group and 3 vs. 1 per counterbalancing conditions within groups). Ages ranged from 18 to 22. Five students identified, at least in part, as African American or Black, 8 as East Asian, 15 as Hispanic or Latinx, 2 as Middle Eastern, 2 as Southeast Asian, Indian, or Pakistani, 2 as Pacific Islander, 39 as White or Caucasian, and 2 as other.

Materials

Because noun recall would be cued with one of two different adjectives (negative or positive), the unit for selecting items was an adjective-adjective-noun triad. Triads from Hertel et al. (2017) had been constructed to be relevant to rumination, so from those 32 triads, we initially chose the 10 that could successfully produce autobiographical memories in a pilot study. Examples included *humiliating-intimate-moment, unappreciated-unexpected-favor, embarrassing-healthy-body*. The remaining materials were selected by choosing nouns from the Affective Norms for English Words (ANEW; Bradley & Lang, 1999), constructing associated

score at the end of the (first) experimental session—a score that keeps them within a few points of the original group designation.

triads, and obtaining pilot ratings for the emotional meaning of each adjective-noun pair (as had been done for the earlier materials in a previous experiment). In the last step we obtained ratings for the ease of retrieving a personal memory, given each of the pairs as a cue. From these pilot data, we chose the 24 emotional triads. In addition, 10 neutral adjective-noun pairs were selected to serve as fillers. All materials are available by request.

Sets. Next, we distributed the 24 triads into four sets of 6 for the purpose of counterbalancing the materials with the experimental conditions, each set to be assigned equally often to each combination of adjective valence and practice status. Triads were distributed to the four sets in a manner that closely balanced the sets on the mean valence rating of each pair, the noun valence from ANEW, and the mean rating for ease of memory generation from the pilot data. The sets were also balanced (less closely) on frequency and number of letters of each member of the triad. Then the 6 triads in each set were divided into two subsets, one to be tested at the end of the first session and the other to be tested at delay for the first time.

Order. For the generation phase, we ordered cue pairs in a randomized block design, by assigning one pair from each set, plus a neutral pair, to each of six blocks; two additional neutral pairs began and ended the list. All subsequent tests preserved that relative ordering by maintaining block membership, but items were randomized anew within the blocks.

Procedure

Double-blind procedures meant that the experimenters did not know RRS scores and the students did not know the basis of their recruitment. To discourage further retrieval practice, the cover story for the delay focused on our need for more memories, but given the length of time that it takes to recall personal memories, we would be requesting more later, and they could be reported

on line, because the procedures have been learned. The procedure for the in-lab session is depicted in Figure 1.

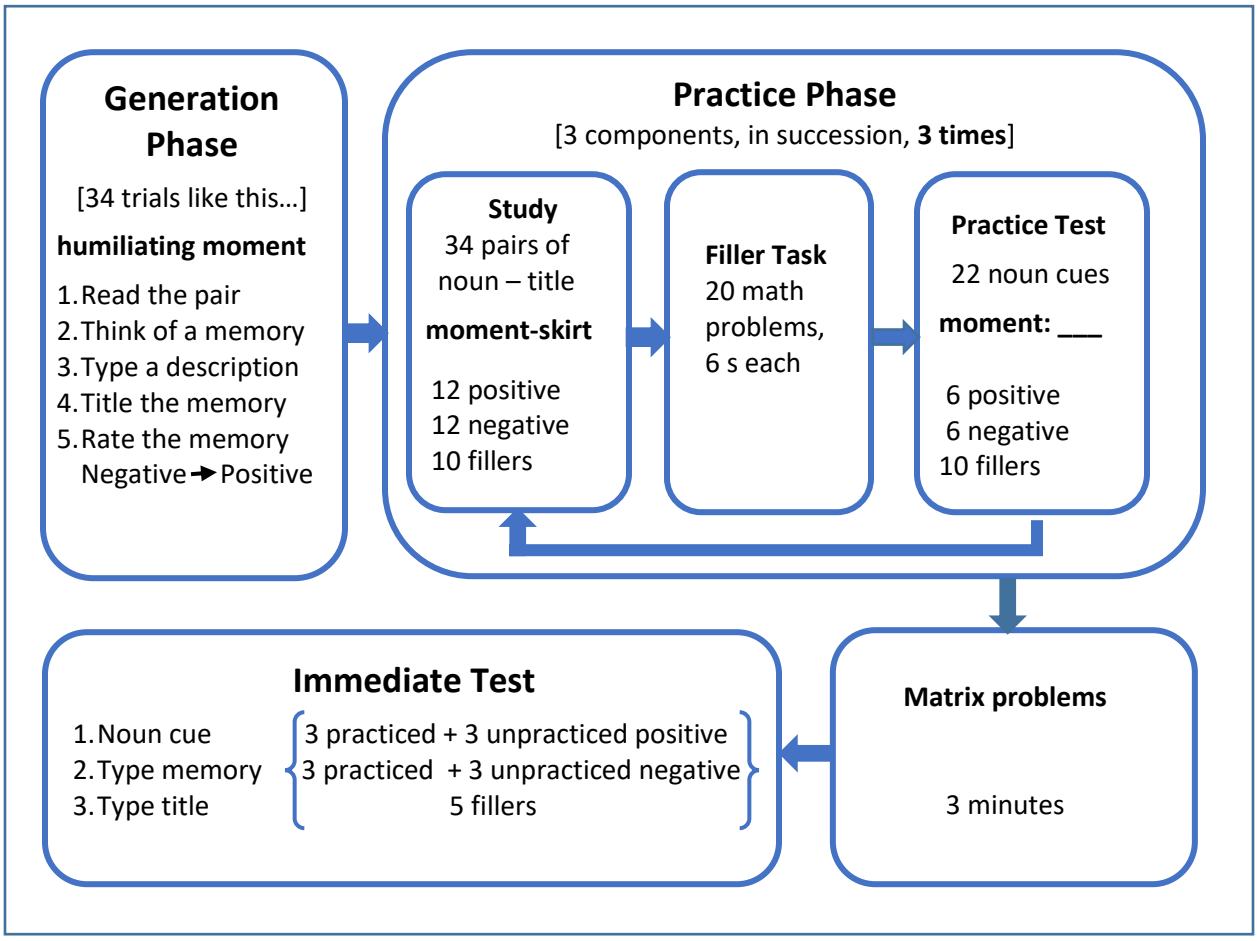

Figure 1. Schematic representation of the segments of procedure in the in-lab session.

Generation phase. Instructions for the generation phase emphasized the importance of recalling memories of events that occurred within a single day and at a particular place (specific memories) at any time prior to one month before the session. The purpose of these specificity instructions was to encourage more detailed memories, not to investigate overgeneral memory; encouraging details made it more likely that the many memories generated would be distinct

from each other—more distinct than they might have been without such instructions.[3] We provided examples of specific and nonspecific memories and told participants that the event could be important or trivial but not extremely unusual ("like what happened to you if you were in one of the recent hurricanes"). We asked participants to avoid reporting memories depicting non-events, such as a general attitude (e.g., "I love Christmas because we travel" when cued by *enjoyable travel*).

At the start of each trial, participants read the cueing pair aloud as they began to recruit an appropriate memory; when the memory came to mind they pressed the space bar and then typed a one-sentence description of the memory. We assured them that the memories would remain completely anonymous, with their identity concealed from the experimenter and everyone who viewed the memories subsequently. After typing the description they pressed the enter key to produce a new screen requesting a title for the memory. We cautioned them to generate a unique title with respect to the other memories they reported in the session, because the title would represent the memory throughout the session. Participants typed the title and reported it aloud, so that the experimenter could record it (for study during practice). Experimenters requested a new title when a previous title was repeated. Subsequent to title report, participants pressed the enter key to reveal a rating scale. The scale was a series of evenly spaced numerals that began with "extremely negative = 1" and ended with "9 = extremely positive." Participants typed a number to represent the memory, and then a new trial began. Before the start of experimental trials, the experimenter walked through a practice trial, answered

[3] There were no significant differences in the percentages of specific memories generated, according to any factor in the design, all *p*s > .28, all $\eta_p^2 < .02$; overall $M = 85.5$, 95% CI [81.3, 89.9]. Out of the recalled memories on the final test, the same pattern obtained, all *p*s > . 47 all $\eta_p^2 < .02$; overall $M = 76.3$, 95% CI [70.3, 82.4]. The apparent reduction likely resulted from the on-line nature of the test (i.e. no experimenter stressing specificity).

questions, made corrections, and then moved behind a screen to record titles. Across all trials, cueing pairs were displayed for 6 s, and all response segments were self-paced.

Practice phase. After a short rest, the next phase consisted of three tasks, presented three times in succession (see Figure 1): guided study of the cueing nouns paired with the participant-generated titles, math problems, and a short memory test. In guided study, the experimenter handed the participants a sheet of paper containing each cueing noun and title and asked them to study each pair for 2 s before moving down the page to the next pair at the sound of a beep (turning the page half way). We asked participants to distribute their study time evenly, without skipping or returning to previous lines. In the math task, 20 simple addition and multiplication problems were each presented on the monitor for 6 s and followed by a screen for typing the answer. Finally, on the practice test, the 12 cueing nouns from the sets identified for practice, plus the 10 nouns from the neutral pairs, were each presented for 4 s. We stressed the importance of thinking about the generated memory during the display, which was followed by a screen for typing the corresponding title (or "forget" if retrieval failed). Again, these three tasks were repeated in succession two more times during the practice phase.

Immediate test. Following another brief rest period and then 3 min during which participants tried to solve problems akin to Ravens' Matrices, we administered the immediate test of half of the practiced and half of the unpracticed memories (12 cues from the counterbalanced sets, and 5 neutral fillers). Each trial began with a 4-s display of the cueing noun, followed by a request to type the memory, and then another request to type the title. Following the test, participants completed the RRS and the Beck Depression Inventory (BDI-II; Beck, Steer, & Brown, 1996, with the suicide item removed).

Delayed test. One week after their participation in the lab session, we emailed the students a link to the Qualtrics test, designed to cue all memories and titles with the nouns from the pairs used at generation. We told participants that they would not be able to stop, save, and start again and that they should find a quiet place with minimal distraction where they could complete the test by 9 pm. All noun cues were presented for 4 s; each was followed by a self-paced interval for typing the memory and then another for typing the title. In the last part of each trial, participants rated the emotional value of the memory on the scale that they had also used for generation, but we told them to judge the emotion in the memory as retrieved on the test and not to report a remembered emotion from a week ago. At the end of the test of 34 trials, the students again responded to the RRS and the BDI-II. They received credit for participating only if they had complied with instructions regarding the timing of the delayed test.

Scoring

Descriptive responses on the immediate and delayed tests were scored for accuracy by two scorers, independently and blind to participant group. Because our purpose did not concern memory for details or degree of emotion, the basis for judging accuracy was simply whether the typed description on the test referred to the memory as originally generated to the same cue. Titles produced during practice and on each test were scored as accurate only if any form of the same word(s) produced during generation was typed in response to the same cueing noun. Scorers were trained on the data from several participants before they scored independently. They agreed about 97.9% of immediately tested memories and 96.4% of memories at delay. A third person resolved all differences. In the final sample, over 90% of the errors were nonresponses (*forget*); the remainder were new memories or responses to other cues.

Results

Figure 2 depicts the important outcomes from this experiment. Retrieval practice increased the recall of the initially generated autobiographical memories and their titles a week later. And if the memories were cued by positive words, ruminators experienced a smaller practice effect. These results from the delayed test are presented first; they are followed by reports of recall during practice and on the immediate test. Finally, we describe the emotion ratings provided for the memories at generation and at delayed testing. The important rating outcome was that the cues functioned to produce emotion-consistent memories.

Percentage of Memories Recalled at Delay

The central questions concerned the percentages of generated memories recalled again on the delayed test, at which point recall was cued by the noun member of the pair used to cue generation (e.g., *paper,* not *failing paper* or *praised paper*). The percentages recalled were submitted to a mixed design analysis of variance, with a between-subjects factor for participant group and within-subject factors for practice status (practiced or not), valence of the pair that cued generation (positive or negative), and the status on the immediate test (tested or not). An additional between-subjects factor was included for counterbalancing condition (4 levels for set distribution across the combination of cue-valence/practice-status) in these and all subsequent analyses, for the purpose of removing systematic variance from the error term. The significant effects involving this factor are not reported.

First, the results replicated the finding of a practice main effect, previously obtained for other types of memories; practiced personal memories were better recalled than unpracticed ones ($M_{practiced}$ = 61.40, SE = 2.55, vs. $M_{unpracticed}$ = 43.81, SE = 2.11), $F(1, 56) = 54.19$, MSE = 730.20, $p < .001$ $\eta_p^2 = .49$. Next, we report effects involving the testing status on the immediate test.

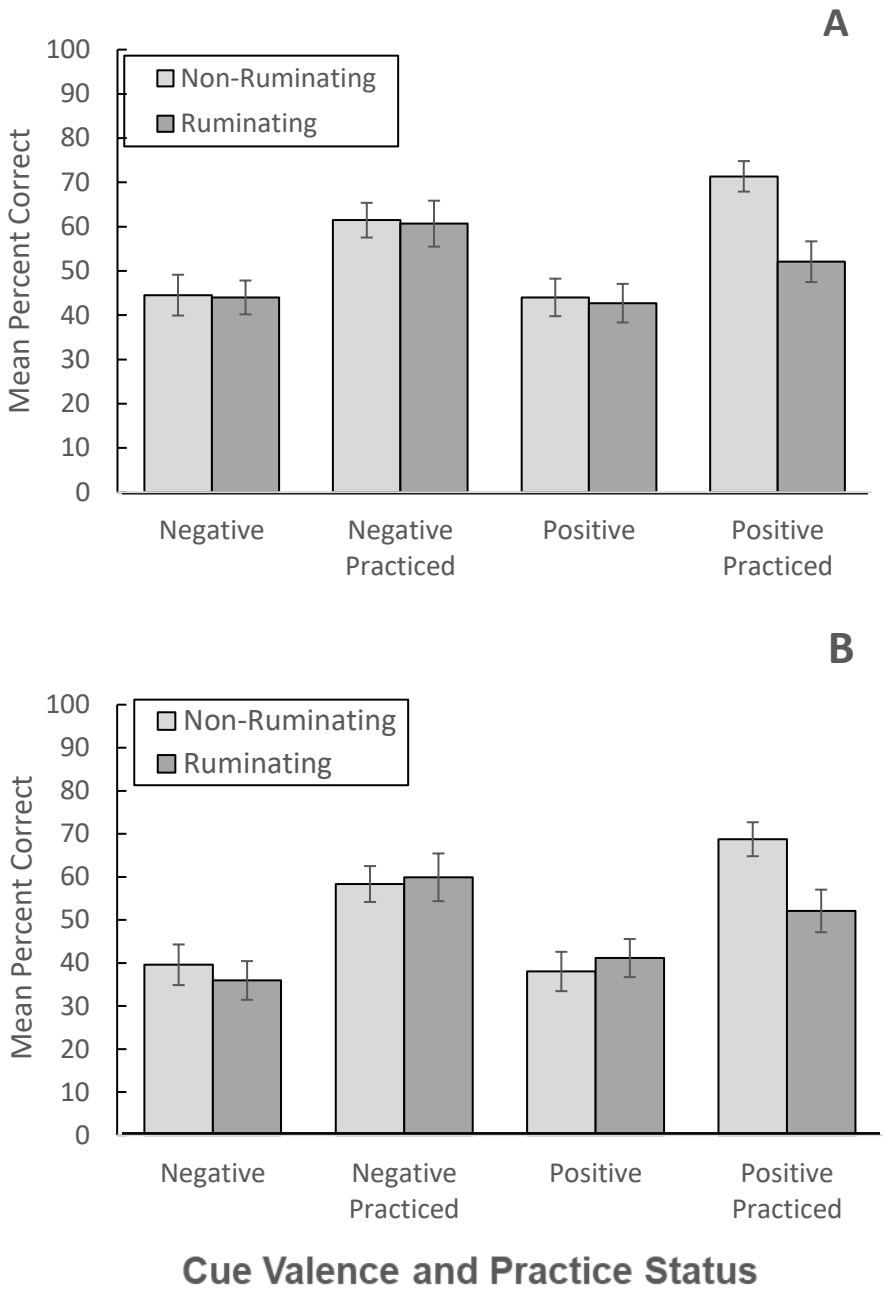

Figure 2. The percentage of memories (A) and titles (B) recalled after a 1-week delay following initial generation. Memories were generated from either negative or positive cueing pairs and then given titles. Half of the memories and titles in each valence condition were practiced three times. Retrieval during all tests was cued with the nouns from the cueing pairs.

Memories tested at the end of the lab session were recalled better at delay than those not test (M_{tested} = 60.87, *SE* = 2.23, vs. M_{not} = 44.33, *SE* = 2.34), $F(1, 56) = 57.83$, *MSE* = 605.82, $p < .001$, $\eta_p^2 = .51$. This outcome clearly reflects the fact that the immediate test provided another (or the only) occasion for retrieval practice, independent of the effect of the three practice trials for half of the memories. Moreover, as can be expected, the effect of immediate testing was smaller for previously practiced memories (M_{tested} = 67.19, *SE* = 3.01, vs. M_{not} = 55.60, *SE* = 2.94) than for those not practiced (M_{tested} = 54.56, *SE* = 2.30, vs. M_{not} = 33.07, *SE* = 2.67), $F(1, 56) = 6.99$, *MSE* = 448.87, $p = .011$, $\eta_p^2 = .11$. No other effect involving immediate testing status reached significance, $p > .35$. And so the effects of primary interest, shown in Figure 2A, can be understood without confounds from immediate testing.

The first of these effects was the interaction of participant group with the valence of the generation cue, $F(1, 56) = 5.02$, *MSE* = 591.92, $p = .029$, $\eta_p^2 = .08$. Regardless of practice status, the two rumination groups recalled similar percentages of memories generated from negative cues (*M* = 53.00, *SE* = 3.16, vs. *M* = 52.34, *SE* = 3.34, for non-ruminators and ruminators, respectively), but ruminators recalled many fewer memories generated from positive cues (*M* = 47.40, *SE* = 3.38, vs. *M* = 57.69, *SE* = 3.29, for non-ruminators). Two additional interactions were marginally significant. One was the interaction of practice status with rumination group, and the other was the three-way interaction of valence, practice, and group, $F(1, 56) = 3.60$, *MSE* = 695.86, $p = .063$, $\eta_p^2 = .06$. We report the statistics for the latter, due to our a priori interest in whether practice effects depend on valence, especially for ruminators.

Follow-up tests subsequent to the three-way interaction examined recall separately for negative and positive cueing pairs. The recall of memories generated to negative cues benefitted substantially from retrieval practice, $F(1, 56) = 21.24$, *MSE* = 425.38, $p < .001$, $\eta_p^2 = .28$, and this

practice effect did not interact with rumination group, $p = .971$. However, as seen in the last four bars in Figure 2A, the simple interaction of practice status with group was significant for memories that had been generated to positive cues, $F(1, 56) = 8.98$, $MSE = 287.65$, $p = .004$, $\eta_p^2 = .14$. The two groups did not differ in recalling unpracticed memories generated to positive cues, $p = .815$. But if these memories had been designated for practice, ruminators recalled fewer, $F(1, 56) = 12.24$, $MSE = 485.86$, $p < .001$, $\eta_p^2 = .18$. [4]

Percentage of Titles Recalled at Delay

In preview, the pattern of effects on delayed recall of titles was strikingly similar to the pattern for memories and suggests that if a memory was recalled, so was its title (although titles were recalled somewhat less frequently in each condition). Quite obviously, practiced titles were recalled better than unpracticed titles on the delayed test, $F(1, 56) = 67.99$, $MSE = 838.03$, $p < .001$ $\eta_p^2 = .55$ (see Figure 2b). Regardless of the other factors, titles tested at the end of the lab session were recalled better a week later than those not tested ($M = 57.0$, $SE = 2.59$, vs. $M = 41.4$, $SE = 2.62$, respectively), $F(1, 56) = 39.85$, $MSE = 84.92$, $p < .001$, $\eta_p^2 = .42$. Moreover, as we found with memory recall, the status of prior testing significantly interacted with the factor for practice, $F(1, 56) = 8.48$, $MSE = 541.64$, $p = .005$, $\eta_p^2 = .13$. Practice produced a greater advantage for titles omitted from the immediate test ($M_{practiced} = 54.95$, $SE = 3.24$, vs. $M_{not} = 27.86$, $SE = 3.04$), compared to immediately tested titles ($M_{practiced} = 64.59$, $SE = 3.16$, vs. $M_{not} = 49.48$, $SE = 2.87$). Otherwise, the status of immediate testing did not interact with other factors, $p > .31$.

[4] The practice effect on positive-memory recall by ruminators was marginally significant, $F(1, 28) = 4.09$, $MSE = 343.94$, $p = .053$, $\eta_p^2 = .13$

In a pattern similar to the pattern for memory recall, the practice effect was qualified by the significant three-way interaction of cue valence, practice status, and rumination group, $F(1, 56) = 7.70$, $MSE = 648.37$, $p = .007$, $\eta_p^2 = .12$. (All other effects in the full design were nonsignificant, $p > .22$.) There were clear effects of title practice when the corresponding memories had been generated to negative cueing pairs, $F(1, 56) = 33.34$, $MSE = 437.87$, $p < .001$, $\eta_p^2 = .37$; and that "negative" practice effect did not interact with group, $p = .485$. And although the "positive" practice effect was even larger, $F(1, 56) = 45.50$, $MSE = 305.33$, $p < .001$, $\eta_p^2 = .45$, it depended on the rumination status of the participant, as shown by the significant simple interaction, $F(1, 56) = 10.25$, $MSE = 305.33$, $p = .002$, $\eta_p^2 = .16$. Another way to understand recall of titles chosen for memories generated to positive cues is to see that the two participant groups recalled unpracticed titles at similar levels, $p = .613$, but the ruminators recalled many fewer practiced titles, $F(1, 56) = 8.01$, $MSE = 554.61$, $p = .006$, $\eta_p^2 = .12$.

Percentage of Titles Retrieved During Practice

We next report percentages of titles correctly recalled during the practice trials. The analysis of variance included a between-subjects factor for participant group (and another for counterbalancing condition) and within-subjects factors for trial number (1 – 3) and the valence of the cue during generation. Data were missing for one non-ruminative participant.

Participants improved with practice and feedback across the three trials ($M = 75.8$, 81.8, and 87.8), $F(2, 109.82) = 49.42$, $MSE = 94.02$, $p < .001$, $\eta_p^2 = .47$ (using the Huynh-Feldt Epsilon to adjust for the violation of sphericity). Relevant to possible evidence for a ruminative bias during practice, all other main effects and interactions were small and nonsignificant. The largest was the interaction of valence with rumination group, $F(1, 55) = 1.73$, $MSE = 450.60$, $p = .193$, $\eta_p^2 = .03$. Descriptively, ruminators recalled somewhat fewer positive titles ($M = 79.5$ vs. 83.7

for non-ruminators), whereas they recalled similar percentages of negative titles (*M* = 82.3 vs. 81.7, respectively).

Percentage of Memories and Titles Recalled, End of Session

With a mixed design identical to the analysis of delayed recall, the percentages of memories and titles recalled at the end of the lab session were separately submitted to an analysis of variance. The percentages of memories recalled at the end of the session did not differ according to any of the factors or their interactions (overall *M* = 82.5, *SE* = 2.06), $p > .31$. Title recall was superior if the titles had been practiced (*M* = 83.7, *SE* = 2.46) than if they had not (*M* = 76.6, *SE* = 3.24), $F(1, 56) = 6.18$, *MSE* = 517.54, $p = .015$, $\eta_p^2 = .10$. All other effects were nonsignificant, $p > .34$.

Emotion Ratings

To investigate whether the emotional quality of the memories corresponded to the valence of the cueing pairs, we evaluated the students' emotion ratings during generation, and to investigate whether the ratings of the recalled memories was affected by practice, the same analysis was conducted for delayed accurate recall. The mean rating across items that qualified in each cell of the design was submitted to separate analyses of variance at generation and delayed testing. Significant differences according to the valence of the cueing pair occurred at each measurement point. At generation, $F(1, 56) = 890.73$, *MSE* = 0.870, $p < .001$, $\eta_p^2 = .94$. At delay, $F(1, 51) = 255.45$, *MSE* = 2.21, $p < .001$, $\eta_p^2 = .83$. At generation, ruminators gave lower (more negative) ratings (*M* = 4.9, *SE* = .089 vs. non-ruminators *M* = 5.2, *SE* = 089), $F(1, 56) = 4.29$, *MSE* = 1.01, $p = .043$, $\eta_p^2 = .07$. However, this group difference disappeared at delay, $p = .193$. All other effects were nonsignificant, $p > .133$. The means at each point are reported in

Table 1

Means (Standard Errors) of Participants' Scores and Rated Emotional Value of Memories

	Generation		Delayed test	
	Non-ruminators	Ruminators	Non-ruminators[a]	Ruminators[b,c]
Emotion ratings				
Negative cueing pairs	3.4 (.14)	3.1 (.14)	3.5 (.18)	3.0 (.17)
Positive cueing pairs	6.9 (.10)	6.6 (.10)	6.4 (.18)	6.4 (.18)
Questionnaire scores				
RRS	30.8 (1.00)	64.2 (1.44)	29.0 (0.98)	58.0 (1.85)
BDI-II	6.7 (0.95)	21.3 (1.81)	6.0 (1.44)	20.9 (1.46)

Note. $n = 32$, except in columns with superscripts; [a] $n = 29$, [b] $n = 30$ for ratings, [c] $n = 31$ for questionnaire measures. Analyses were not performed on the RRS and BDI-II scores, due to non-independent sampling.

Table 1,[5] along with the mean scores on the RRS and BDI-II at the end of the lab session and a week later.

Discussion

Autobiographical memories are more likely to come to mind when cued if they have been previously practiced. More specifically, our demonstration of this practice effect applies to intentional recall that benefits from multiple prior repetitions of the same intentional act.

[5]An additional analysis of emotion ratings that included a factor for measurement point did reveal a significant main effect for session, $F(1, 51) = 7.71$, $MSE = 0.680$, $p = .008$, $\eta_p^2 = .13$, and a nonsignificant trend for an interaction of valence with session, $F(1, 51) = 3.66$, $MSE = 0.893$, $p = .061$, $\eta_p^2 = .07$. On average, ratings were 0.21 points lower at delay—0.05 points lower for negative memories and 0.37 points lower for positive memories. Interpretation must take into account the confound that means were calculated across varying numbers of memories, due both to forgetting and to the uncontrolled conditions for the delayed test.

Similarly, ruminators repeatedly and sometimes intentionally focus on past events and thereby produce a readily accessible memory for future episodes of rumination.

The practice effect in our results might have been aided by the use of a descriptor—we called it a title—and that feature is also not unlike common experience. Our participants frequently chose titles that were names of people or places, or topics that might naturally remind them of the memory if they occurred in the course of a typical day. For example: To the generation cue *thoughtless friend*, one of our participants remembered that he had "made a derogatory remark about 'life choices' to a pansexual friend, and later realized how callous what [he] said was and apologized." He titled the memory "Insult" and remembered the title and the memory a week later, even though it had not been assigned for practice.

As the example possibly illustrates, ruminators sometimes focus on past events in order to create a new understanding or a better resolution of conflict inherent in at least the interpretation of the episode, if not its reality. They believe that such acts of remembering might help them feel better in the long run, even though rumination often worsens their moods (Takano, Van Grieken, & Raes, 2019). And each time this memory comes to mind, again it becomes even more available for future fruitless access. Such remembering is self-perpetuating.

Our results also suggest that self-perpetuating memories are equally applicable to people who do not ruminate…if they would happen to initiate the process of practicing recall of specific memories, negative or positive. And maybe this is a topic relevant, for example, to late-life recollection: Intentional reminiscence perpetuates itself and contributes to the positive memory bias in aging populations (see Mather, 2006). But the point for our purpose is that ruminators are the people who actually do initiate the practice of repeated recall, and compared to non-

ruminators, instructions to practice recalling positive events was unfortunately less productive for future remembering.

Our secondary finding—that the practice of positive memories was not as beneficial for ruminators—was inconsistent with the hope that positive practice might overcome ruminative recall (see Hertel et al., 2017). The term rumination refers to repetitive thought patterns that typify depressed people, and there are similar asymmetric patterns of recall in the literature on depression and memory. (The two samples in this experiment separate on BDI-II scores, with a few exceptions.) Among the several findings that suggest a depression-related impairment in the recall of positive memories is evidence that happy memories are harder to retrieve (Rottenberg, Hildner, & Gotlib, 2006) and that fewer memories are produced to positive cueing words (Young, Erickson, & Drevets, 2012). Similarly, compared to others, people scoring high on an on-line depression measure less effectively revised negative interpretations after reading disconfirmatory positive information (Everaert, Bronstein, Cannon, & Joormann, 2018). On the initially hopeful side, we might consider recent evidence that a positive bias in recalling target *words*, established through retrieval practice (training), transferred to dysphoric students' autobiographical recall, but transfer occurred only if the students had experienced a positive bias before training (Vrijsen et al., 2019).

Both Vrijsen et al. (2019) and Hertel et al. (2017) successfully trained delayed recall of positive words by ruminators, yet the current experiment was less successful in training recall of positive personal memories. How can we understand this difference? Positive *personal memories* sometimes are accompanied by increasingly sad moods for people categorized as depressed (e.g., Joormann, Siemer, & Gotlib, 2007), possibly because the positive memories incur ruminative comparisons with their current state, and positive *words* lack the connections for cueing such

comparisons, especially Swahili translations (Vrijsen et al., 2019). This difference therefore suggests that positive personal memories might have been attended less well (e.g., less elaboratively) either during practice trials or subsequently in the week prior to the final test. In addressing this speculation, we found little support for the idea that ruminators recalled fewer positive titles during practice, and we have no evidence concerning the extent to which they brought the full memories to mind during practice or later in the week. We do, however, know that positive titles and memories were recalled equally well by ruminators and nonruminators at the end of the lab session. Therefore, a sensible way to understand that autobiographical recall might show different practice effects, compared, to word recall, is to suggest that the positive memories were entertained less often during the delay before the final test, or not fully enough during practice to sustain recall beyond the session.[6]

Before we settle on the idea that positive autobiographical recall is doomed to minimal practice effects for ruminators, another procedural difference should be considered. In the two experiments on retrieval practice with words, participants practiced words from just one condition of emotional valence. In the experiment by Hertel et al. (2017) they practiced recalling nouns cued by *either* positive or negative adjectives (and neutral ones), and in the experiment by Vrijsen et al. (2019) they practiced English translations of positive *or* neutral Swahili words. In the current experiment, participants practiced positive *and* negative memories. A reasonable suggestion, therefore, is that the coincidental practice of negative memories might hold attention, even during the positive trials. There is some evidence to suggest that ruminators have trouble

[6] Readers might notice that ruminators and nonruminators recalled unpracticed positive and negative memories equally well. We assume that that this baseline level of recall reflects the probability that the test cues work similarly to their cuing tendency during the generation phase, as well as the similar tendencies to forget the episodic context of having produced them a week earlier.

switching attention away from active negative content (Joormann & Gotlib, 2008; Koster, De Lissnyder, & De Raedt, 2013).

Before summarizing, we offer yet another consideration regarding these results. Our experimental model of memory-inbued rumination assumes deliberate attempts to retrieve memories during practice, whereas the real-world version of remembering during rumination is more likely involuntary. Follow-up experiments might try to encourage involuntary retrieval (by presenting the cues multiple times in a different task, for example). As a starting place, however, we thought it was important to ensure retrieval by deliberate invitation, given the limitation of experimental sessions and the associated difficulties involved in cueing involuntary memories over longer stretches of time.(see Berntsen, 1998). Extrapolating from other comparisons of voluntary and involuntary autobiographical memories leads us to imagine that the valence difference we obtained for voluntary practice effects would be even larger for involuntary practice effects (Schlagman & Kvavilashvili, 2008; cf. Watson, Berntsen, Kuyken, & Watson, 2012).

A final note concerns our difficulties in data collection, apparent by examining the amount of data we set aside and replaced (see our description of participants). One major reason for the need to replace was the failure to complete the final test, even after reminders. We can speculate that these students might have reconsidered their need for experiment credit or their willingness to spend the time and effort. We also lost participants who believed our cover story about new memories and ignored the test instructions; clearly we should have created a different cover story. The largest reason for loss—change in rumination status by the end of the lab session—is difficult to understand. Some participants (often dysphoric) were scheduled as long as a month after screening, mainly because they were difficult to recruit. On the other hand, even

though the RRS is understood to be a measure of trait rumination, some RRS changes occurred for participants who participated soon after screening. Are such changes indications of exaggeration during class screening, or are they reflections of either regression to the mean or reticence to reveal their thoughts during individual administration? Not knowing why they occurred, we chose a relatively conservative approach in requiring the session scores to be within 3 points of the cut-off at screening. Our sense, therefore, is that the final samples do represent ruminators and nonruminators.

In conclusion, we emphasize the finding that retrieval practice serves to perpetuate retrieval of autobiographical memories. From a clinical point of view, such an outcome is potentially good or bad for ruminators: Bad if it perpetuates ruminative recall; good if it works to encourage happy memories (and ruminators did indeed produce a marginally significant practice effect for positive memories). We can speculate, for example, that the clinical technique of discussing negative experience is accompanied by the danger of perpetuating its recall. We can also suggest that if happy memories are practiced in the context of also practicing ruminative memories, the good outcomes will be less accessible for ruminators. Future research might more closely replicate the procedures used by Hertel et al. (2017) and ask ruminators to practice only positive personal memories. We hope we do not seem too pessimistic in suggesting that asking ruminators to practice only happy personal memories outside the laboratory or therapy session is at best a slow path to change.

Author Contributions

PTH and PH designed the experiment and constructed materials. Data collection and scoring were performed by PH and PS, who also wrote previous versions of the paper. PTH analyzed the data and wrote the final version, and all authors approved it.

Conflict of Interest

The authors declare no conflicts of interest in the preparation of this research, data collection, or publication.

References

Beck, A. T., Steer R. A., & Brown, G. K. (1996). Manual for the Beck Depression Inventory (2nd Ed.). San Antonio, TX: Psychological Corporation.

Berntsen, D. (1998). Involuntary autobiographical memories in and outside the laboratory: How different are they from voluntary autobiographical memories? *Memory, 6,* 113-141.

Bjork, R. A. (1975) Retrieval as a memory modifier: An interpretation of negative recency and related phenomena. In R. Solso (Ed.), *Information processing and cognition: The Loyola Symposium* (pp. 123-144), Earlbaum: Hillsdale, NJ.

Cohen, N., Mor, N., & Henik, A. (2015). Linking executive control and emotional response: A training procedure to reduce rumination. *Clinical Psychological Science, 3,* 15-25.

Daches, S., Mor, N., & Hertel, P. (2019). Training to inhibit negative content affects memory and rumination. *Cognitive Therapy and Research*, doi.org/10.1007/s10608-019-10023-0

Everaert, J., Bronstein, M. V., Cannon, T. D., & Joormann, J. (2018). Looking throught tinted glasses: Depression and social anxiety are related to both interpretation biases and inflexible negative interpretations. *Clinical Psychological Science, 6,* 517-528.

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., Mor, N., Ferrari, C., Hunt, O., & Agrawal, N. (2014). Looking on the dark side rumination and cognitive-bias modification. *Clinical Psychological Science, 2,* 714–726.

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, 6,* 872–881.

James, W. (1890). *Habit* (p. 24). New York: H. Holt.

Joormann, J., & Gotlib, I. H. (2008). Updating the contents of working memory in depression: Interference from irrelevant negative material. *Journal of Abnormal Psychology, 117*, 182–192.

Joormann, J., Siemer, M., & Gotlib, I. H. (2007). Mood regulation in depression: Differential effects of distraction and recall of happy memories on sad mood. *Journal of Abnormal Psychology, 116,* 484-490.

Joormann, J., & Tran, T. B. (2009) Rumination and intentional forgetting of emotional material. *Cognition and Emotion. 23*, 1233–1246.

Koster, E. H. W., De Lissnyder, E., & De Raedt, R. (2013). Rumination is characterized by valence-specific impairments in switching of attention. *Acta Psychologica, 144*, 563-570.

Lyubomirsky, S., Caldwell, N. D., & Nolen-Hoeksema, S. (1998). Effects of ruminative and distracting responses to depressed mood on retrieval of autobiographical memories. *Journal of Personality and Social Psychology, 75*, 166–177.

Mather, M. (2006). Why memories may become more positive as people age. In B. Uttl, N. Ohta, & A. L. Siegenthaler (Eds.), *Memory and emotion: Interdisciplinary perspectives* (pp. 135-158). Malden, : Blackwell Publishing.http://dx.doi.org/10.1002/9780470756232.ch7

Morris, C. D., Bransford, J. D., & Franks, J. J. (1977). Levels of processing versus transfer appropriate processing. *Journal of Verbal Learning and Verbal Behavior, 16*, 519-533.

Nolen-Hoeksema, S., Wisco, B. E., & Lyubomirsky, S. (2008). Rethinking rumination. *Perspectives on Psychological Science, 3,* 400-424.

Roediger, H. L., III, & Butler, A. C. (2011). The critical role of retrieval practice in long-term retention. *Trends in Cognitive Sciences, 15*, 20-27.

Rottenberg, J., Hildner, J. C., & Gotlib, I. H. (2006). Idiographic autobiographical memories in major depressive disorder. *Cognition and Emotion, 20,* 114-128.

Schlagman, S., & Kvavilashvili, L. (2008). Involuntary autobiographical memories in and outside the laboratory: How different are they from voluntary autobiographical memories? *Memory & Cognition, 36,* 920-932.

Takano, K., Van Grieken, J., & Raes, F. (2019). Difficulty in updating positive beliefs about negative cognition is associated with increased depressed mood. *Journal of Behavior Therapy and Experimental Psychiatry, 64*, 22-30

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

Vrijsen, J. N., Dainer-Best, J., Witcraft, S. M., Papini. S., Hertel, P., Beevers, C. P., … Smits, J. A. J. (2019) Effect of cognitive bias modification-memory on depressive symptoms and autobiographical memory bias: two independent studies in high-ruminating and dysphoric samples, *Cognition and Emotion, 33,* 288-304.

Vrijsen, J.N., Hertel, P. T., & Becker, E.S. (2016). Practicing emotionally biased retrieval affects mood and establishes biased recall a week later. *Cognitive Therapy and Research, 40,* 764–773.

Watson, L. A., Berntsen, D., Kuyken, W., & Watson, E. R. (2012). The characteristics of involuntary and voluntary autobiographical memories in depressed and never depressed individuals. *Consciousness and Cognition, 21,* 1382-1392.

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

Young, K. D., Erickson, K., & Drevets, W. C. (2012). Match between cue and memory valence during autobiographical memory recall in depression. *Psychological Reports: Mental & Physical Health, 111,* 129-148.

Zetsche, U., Burkner, P.-C., & Schulze, L. (2018). Shedding light on the association between repetitive negative thinking and deficits in cognitive control—A meta-analysis. *Clinical Psychology Review, 63,* 56-65.