11-2005

Remembering the Good, Forgetting the Bad: Intentional Forgetting of Emotional Material in Depression

J. Joormann

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

F. Brozovich

I. H. Gotlib

Publication Details
Journal of Abnormal Psychology

Journal of Abnormal Psychology
2005, Vol. 114, No. 4, 640–648

0021-843X/05/$12.00 DOI: 10.1037/0021-843X.114.4.640

Remembering the Good, Forgetting the Bad: Intentional Forgetting of Emotional Material in Depression

Jutta Joormann
Stanford University

Paula T. Hertel
Trinity University

Faith Brozovich and Ian H. Gotlib
Stanford University

The authors examined intentional forgetting of negative material in depression. Participants were instructed to not think about emotional nouns that they had learned to associate with a neutral cue word. The authors provided participants with multiple occasions to suppress the unwanted words. Overall, depressed participants successfully forgot negative words. Moreover, the authors obtained a clear practice effect. However, forgetting came at a cost: Compared with the nondepressed participants and with the depressed participants who were instructed to forget positive words, depressed participants who were instructed to forget negative words showed significantly worse recall of the baseline words. These results indicate that training depressed individuals in intentional forgetting could prove to be an effective strategy to counteract automatic ruminative tendencies and mood-congruent biases.

Keywords: intentional forgetting, rumination, depression memory, bias

One of the more troubling aspects of depression is the frequent occurrence of unintentional and often uncontrollable negative thoughts and memories. Not only are these negative cognitions a debilitating symptom of depression, but they are also related to both the maintenance of depressive episodes and the likelihood of recurrence (e.g., Nolen-Hoeksema, 2000; Nolen-Hoeksema & Larson, 1999; Roberts, Gilboa, & Gotlib, 1998). Therefore, gaining a better understanding of conditions that might decrease the occurrence of negative thoughts and memories in depressed individuals is an important first step in the development of effective methods of remediation.

One of the most consistent and robust findings of the numerous investigations of memory processes associated with depression is that depressed participants are characterized by a mood-congruent recall bias, recalling more negative than positive words on explicit memory tasks (Matt, Vazquez, & Campbell, 1992). In these studies, depressed individuals are typically asked to learn a list of words or to judge words in terms of self-relevance and then to recall these words. In attempting to understand the reason for the depression-associated preferential recall of negative material in these studies, cognitive theorists have suggested that depressed individuals engage in strategic elaboration of negative information (Williams, Watts, MacLeod, & Mathews, 1997). Increased elaboration of negative material, either during or following the initial encoding phase, increases the subsequent accessibility and ease of retrieval of negative concepts. Empirical support for this explanation comes from a related line of research demonstrating that depression is associated with a tendency to respond to negative life events and negative mood states by ruminating about the event or mood (Lyubomirsky & Nolen-Hoeksema, 1993; Nolen-Hoeksema, 2000; Nolen-Hoeksema, Morrow, & Fredrickson, 1993). Like elaboration, rumination makes negative events and memories more accessible and more likely to be easily retrieved, resulting in a vicious circle of rumination, mood-congruent recall, and depressed mood state (Lyubomirsky, Caldwell, & Nolen-Hoeksema, 1998).

Although depressed and nondepressed individuals appear to differ in their habitual processing of negative material, it is not clear that they would differ in their ability to intentionally forget this material. Indeed, in contrast to nondepressed persons, depressed individuals typically make little effort to suppress and forget negative thoughts (Lyubomirsky & Nolen-Hoeksema, 1993). It is possible, however, that with guided and repeated attempts to forget negative material, depressed persons could reduce elaborative processing and, consequently, negative recall. It is important to note that studies examining free recall and recognition by depressed persons of experimentally presented valenced stimuli are assessing passive forgetting; participants in these experiments are not instructed to forget or suppress material. However, forgetting can also take the form of an active process that is deliberately instigated (Levy & Anderson, 2002). In addition, intentional forgetting, when applied to negative material, could develop into a powerful emotion-regulation strategy. Indeed, studies of emotion regulation in older participants have consistently reported that forgetting the bad times and remembering the good times are strongly associated with well being and overall life

Jutta Joormann, Faith Brozovich, and Ian H. Gotlib, Department of Psychology, Stanford University; Paula T. Hertel, Department of Psychology, Trinity University, San Antonio, TX.

This research was supported by a grant from the German Research Foundation (DFG JO/399-1 and JO/399-2) to Jutta Joormann and Grant MH59259 from the National Institute of Mental Health to Ian H. Gotlib.

Correspondence concerning this article should be addressed to Jutta Joormann, Department of Psychology, Stanford University, Stanford, CA 94305. E-mail: joormann@psych.stanford.edu

satisfaction (Charles, Mather, & Carstensen, 2003; Kennedy, Mather, & Carstensen, 2004). Few studies have investigated the consequences of intentional forgetting on the recall of negative material in depression. We designed the present investigation to examine whether training depressed participants in suppression and giving them explicit instructions to forget facilitated their forgetting of negative material; if so, the results of this study might contribute ultimately to the development of an effective method of remediation.

A small number of studies have used a "directed forgetting" task to examine individual differences in instructed forgetting (e.g., Korfine & Hooley, 2000; Tolin, Hamlin, & Foa, 2002). Of these, in only one study did the authors examine intentional forgetting by participants diagnosed with major depressive disorder. Power, Dalgleish, Claudio, Tata, and Kentish (2000) used a directed forgetting paradigm in which, halfway through the learning phase, depressed and nondepressed participants were instructed to forget the words they had learned so far. When participants were tested on their final recall for words from both halves of the list, only the depressed participants showed better memory for the to-be-forgotten negative than for the to-be-forgotten positive words in the first half of the list. Although these findings suggest that depression is associated with impaired intentional forgetting of negative material, it is important to note that the degree to which this paradigm represents intentional forgetting is questionable. Participants were instructed to forget, but the instruction occurred just once, in the absence of cues to remind them about the to-be-forgotten words, and just prior to the learning of new materials. In short, the forgetting processes used and assessed in the directed forgetting paradigm are arguably more passive than active. (For a related analysis of directed forgetting, see Sahakyan & Kelley, 2002.)

In an important series of studies using primarily nonemotional material in unselected samples, Anderson and Green (2001) demonstrated that forgetting instructions and training in suppression of unwanted thoughts can have enduring consequences for subsequent recall. In contrast to directed forgetting paradigms, Anderson and Green's "think–no-think" procedure required participants to practice suppression when faced with multiple occurrences of the cue that they had learned to associate with the response to be suppressed. This think–no-think procedure clearly involves active suppression of material. Anderson and Green found that when people consistently prevent memories from entering awareness, the subsequent deliberate recall of the rejected memories becomes increasingly difficult. Participants in these studies first learned unrelated word pairs to criterion. In the next phase, the first member of some pairs were presented as cues for recalling the second member (respond condition), and the first member of other pairs served as a cue to stop the second member from coming to mind (suppress condition). Last, after a varied number of trials in making responses to some cues and suppressing responses to others, participants were asked to recall all response words. Using this think–no-think design, Anderson and Green showed that the level of recalling previously suppressed words dropped below the baseline recall of words not cued in the second phase. Moreover, the greater the number of suppression trials was, the fewer the words that were recalled on the final test. Therefore, thought suppression appears to make the subsequent recall of suppressed thoughts or memories more difficult, at least for the relatively neutral words used in Anderson and Green's studies. It is possible that different results would be obtained with emotional materials and that such results would depend, in turn, on whether the research participants are depressed.

Using a variant of Anderson and Green's (2001) think–no-think paradigm, Hertel and Gerstle (2003) had dysphoric and nondysphoric students learn initially neutral nouns paired with positive or negative adjectival cues (e.g., *gloomy cottage* vs. *splendid cottage*; *slashed skin* vs. *smooth skin*). During the learning phase, participants were instructed to form a self-referential mental image for each pair. In the think–no-think phase, the valence of the cue was counterbalanced with the instruction to respond with or suppress the corresponding noun. On the final test of cued recall, Hertel and Gerstle found that dysphoric students recalled more previously suppressed nouns than did controls, regardless of the valence of the cue. Although Hertel and Gerstle had predicted that forgetting would be a function of the valence of the cue, in retrospect the absence of valence effects is not surprising. Studies using standard memory paradigms with dysphoric students have yielded inconsistent evidence for mood-related biases; instead, dysphoric individuals tend to show evenhanded recall (Matt et al., 1992). Clinically depressed participants, in contrast, typically produce clear evidence of memory bias in these standard paradigms and, therefore, might show such a bias in intentional forgetting.

The goal of the present study was to use Anderson and Green's (2001) think–no-think paradigm to investigate intentional forgetting of emotional words in carefully diagnosed, clinically depressed participants. We were interested in examining whether, after learning a set of cue–response pairs to criterion, subsequent practice in suppressing the valenced response words would have consequences for later recall. The main question was whether we could reduce or eliminate the consistently reported tendency for depressed participants to recall negative stimuli by giving them instructions to forget and by giving them practice in suppression. Specifically, we predicted that depressed participants instructed to suppress thinking about negative words would not show mood-congruent recall, but would show reduced retrieval of the negative words on subsequent testing. If this result is obtained, training in intentional forgetting could prove to be a useful emotion-regulation strategy for individuals suffering from depression.

This is the first study using the think–no-think paradigm with neutral cues and emotional target words to investigate intentional forgetting for emotional words. The first purpose of our study, therefore, was to replicate Anderson and Green's (2001) finding of below-baseline recall of the target words as a function of suppression training in a group of nondepressed participants. In this context, our first hypothesis was that nondepressed participants would exhibit a linear decline in recall rates of the target words as a function of the number of times they practiced suppressing these words in the presence of the cue words, and they would exhibit a linear increase in recall rates of the target words as a function of the number of times they practiced responding with these words in the presence of the cue words. We also expected that the nondepressed participants would not exhibit differential recall rates as a function of the valence of the target words. Our primary interest, however, was in the effects of the suppression training on depressed participants who were required to respond to neutral cues by recalling or suppressing negative words that had been paired with these cues. Specifically, in our second hypothesis, we pro-

posed that depressed participants who were trained to suppress negative target words would not exhibit the preferential recall of negative words that has been found consistently in previous research to be associated with depression, but would exhibit below-baseline recall of the negative words. Thus, we predicted that depressed participants would exhibit a linear increase in recall rates of the negative target words as a function of the number of times they practiced responding with these words in the presence of the cue word and a linear decline in recall rates of the negative target words to below-baseline rates of recall as a function of the number of times they practiced suppressing these words in the presence of the cue word.

Method

Overview

The think–no-think task consisted of three separate phases. In the first—the learning phase—participants studied the neutral cue word–valenced response word pairs (e.g., "lamp–pain", "curtain–humor") and were asked to memorize them for later recall. At the end of the learning phase, participants were tested on their knowledge of the cue–response pairs. The cue words (e.g., lamp, curtain) were presented and the participants were asked to recall the associated response words (e.g., pain, humor); feedback was provided. Participants were required to reach the criterion of 50% correctly recalled response words to move on to the second phase. In the second phase—the suppression phase—participants were trained either to suppress the associated words or to respond to the neutral cue words with the associated words multiple times. Cue words that were presented in green on the computer screen indicated that the associated response words should be recalled, and cue words that were presented in red signaled that the participants should prevent the response words from coming to mind. The cues were presented either 0, 2, or 12 times. Thus, some of the cues of the cue–response word pairs that the participants had learned in the first phase of the task were never presented in the suppression phase. Accordingly, participants never practiced responding with or suppressing the associated response words. The recall of these response words on the final test served as the baseline condition. Half of the depressed and half of the control participants were asked to suppress positive response words in response to the neutral cues and to respond with negative response words, whereas the other half of each group were asked to suppress negative response words and to respond with positive response words. Last, in the third phase—the final recall test phase—participants were instructed to recall all response words in response to the cues, regardless of whether they had practiced recalling or suppressing them.

Participants and Design

Procedure for selection. Volunteers responded to online advertisements concerning information-processing research. These online advertisements are published under "Job Opportunities" on a frequently visited, easily accessible, Internet page covering the San Francisco area. In our advertisement, readers were asked whether they were currently experiencing depressed mood for most of the day or had lost pleasure in most of their daily activities and, if so, whether this had lasted for 2 weeks. In addition, readers were told that if this described them, they might be eligible to take part in research examining how people think or process different kinds of situations. We prescreened respondents over the telephone and then invited those who met screening criteria (meant to identify potentially diagnosable depressed individuals and nondepressed controls) to come to the laboratory for a more extensive interview. Trained interviewers administered the Structured Clinical Interview (SCID; First, Spitzer, Gibbon, & Williams, 1996) for the *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition* (*DSM–IV*; American Psychiatric Association, 1994) to all participants during their first session in the study. Participants were included in the depressed group if they met *DSM–IV* criteria for major depressive disorder (MDD). The nondepressed group consisted of individuals with no current diagnosis and no history of any Axis I disorder. Severe head trauma, color blindness, and learning disabilities excluded individuals from the study, as did current panic disorder, current social phobia, psychotic symptoms, bipolar disorder, and alcohol or substance abuse within the past 6 months. Although we did not exclude all comorbid disorders in our MDD group, overall comorbidity rates in our final sample were low. Only 7 of our 36 depressed participants were diagnosed with a comorbid disorder: 3 with current posttraumatic stress disorder (PTSD), 1 with past PTSD, 1 with binge eating disorder, 1 with dysthymia, and 1 with obsessive–compulsive disorder (OCD). Individuals who met the inclusion criteria were scheduled for a second session of "computer tasks," typically conducted within 2 weeks of the interview.

At the end of the first and second sessions, participants completed the Beck Depression Inventory (BDI-II; Beck, Steer, & Brown, 1996). Only the nondepressed participants who scored below 9.0 on the BDI-II at the second session were included in the final sample. The mean BDI score was 1.8 in the control group and 27.5 in the MDD group. Within the depressed and control groups, BDI scores did not differ significantly as a function of assigned experimental condition ($p > .10$ within the depressed group, $p > .50$ within the nondepressed group).

Design and final sample. Ninety-one participants (40 participants who were depressed and 51 nondepressed participants) were administered the think–no-think task. The data from 4 depressed and 15 nondepressed participants were excluded because they did not achieve the 50% criterion during the learning phase. Thus, the final sample consisted of 72 participants (36 participants who were depressed and 36 nondepressed participants; 17 men and 55 women). Participants were randomly assigned to each combination of valence condition (i.e., suppress-positive–respond-negative; suppress-negative–respond-positive) and the six counterbalancing conditions, subject to the constraint of equal cell sizes. The participants in the depressed and nondepressed groups did not differ in age, ($M = 34.2$ years for both groups), or education (69% of depressed and 78% of nondepressed participants were college graduates), $\chi^2[2, N = 72] < 1.0$; both ps $> .05$). In addition, distribution according to gender was similar across conditions. There were 9 male (5 control, 4 MDD) and 27 female (13 control, 14 MDD) participants in the suppress-positive–respond-negative condition and 8 male (5 control, 3 MDD) and 28 female (13 control, 15 MDD) participants in the suppress-negative–respond-positive condition.

Materials

Stimuli. Thirty-six negative, 36 neutral, and 36 positive nouns were selected from the Affective Norms for English Words (ANEW; Bradley & Lang, 1999). Words with affective ratings below 4.0 (on a 9-point scale) were eligible for the negative category, between 4.0 and 6.0 for the neutral category, and above 6.0 for the positive category. (Respective mean ratings were 2.3, 5.4, and 7.5.) Other criteria for the word selection included average word length (negative $M = 6.1$, neutral $M = 5.5$, positive $M = 6.2$), frequency (negative $M = 13.7$, neutral $M = 30.9$, positive $M = 20.9$), and arousal (negative $M = 5.6$, neutral $M = 4.0$, positive $M = 5.4$). The lists of positive and negative words did not differ significantly in word length, $t(70) < 1$; arousal ratings, $t(70) = 1.12$, $p > .05$; or word frequency, $t(70) = 1.70$, $p > .05$. Because ANEW does not provide concreteness ratings, 20 Stanford University students were asked to rate concreteness. We also used the University of South Florida association norms (Nelson, McEvoy, & Schreiber, 1998) to limit the number of forward and backward associations within the positive and negative lists and between the neutral and both emotional lists. Within each valence group the words were organized into six sets containing six words each, balanced on mean word length, frequency, arousal, and concreteness (all

*F*s < 1). (Fourteen neutral filler words, to constitute seven pairs, were selected from the same pool and criteria for use in the learning and suppression phases of the experiment.)

The six sets of neutral words were then paired separately with the six sets of positive words and the six sets of negative words. Using the University of Florida association norms, we ensured that there were no reported forward or backward intrapair associations. Finally, we created two lists, each of which contained three neutral–negative sets and three neutral–positive sets (i.e., the neutral nouns in Sets 1–3 were paired with negative nouns in List 1 and positive nouns in List 2; neutral nouns in Sets 4–6 were paired with positive nouns in List 1 and negative nouns in List 2). Thus, participants learned either negative or positive response words to the same cue word but learned both negative and positive responses overall (e.g., *trunk disaster*, *mushroom pillow* or *trunk breeze*, *mushroom hostage).* Within each list half—positive or negative response words—the three sets were rotated systematically across the three conditions of cue presentation during the suppression phase (0, 2, or 12 cue presentations). Together, this rotation and the list manipulation comprised six conditions for counterbalancing materials with conditions of the independent variables of response valence and number of cue presentations.

Questionnaires. At the end of the session, participants completed the Strategies Questionnaire and the BDI-II. The BDI-II is a reliable and widely used self-report measure of depressive symptoms (Beck et al., 1996). It consists of 21 items that assess the levels of depressive symptoms present in the past 2 weeks. The Strategies Questionnaire was adapted from a measure used by Anderson (personal communication, May 18, 2002) and modified by Hertel and Calcaterra (2005). Participants answered each strategy question on the following scale: 0 (*never*), 1 (*rarely*), 2 (*sometimes*), 3 (*frequently*), and 4 (*very frequently*). Participants rated the extent to which they used each strategy during the suppression phase. The items concerned whether participants (a) first made sure that they knew the response word before not thinking about it, (b) made sure they still remembered the response word after the trial was over, (c) kept repeating the response word silently while not saying it aloud, (d) kept themselves from thinking about the response word by thinking about something else, and (e) kept themselves from thinking about the response word by keeping their minds completely blank.

Procedure

All tasks were implemented with Superlab Pro software (Cedrus Corporation, San Pedro, CA). The procedure was adapted from Hertel and Gerstle (2003). The experimental session lasted 60–90 min, and participants were each paid $25 per hour.

Learning phase. Forty-three word pairs each appeared in black font, centered on a white computer screen for 5 s. Participants were instructed to study the word pairs for a later test. The word pairs appeared in a randomized block order. Each block of six trials contained 1 pair from each of the six sets. Three of the 7 filler pairs appeared at the beginning of the list, 3 at the end, and 1 in the middle. At the end of the learning phase, we administered a feedback test during which participants recalled response words out loud. Cues appeared on the screen in black font for 5,200 ms (or less if the response occurred sooner). Immediately following cue offset, the correct response was displayed in blue font for 2 s. The participants were told to use the blue feedback to reinforce their knowledge of the word pair. If they recalled fewer than 50% of the word pairs correctly, we tested their knowledge again with differently ordered cues. Participants completed up to four feedback tests to learn the word pairs to 50% criterion to advance to the suppression phase. If they failed to reach the 50% criterion after four feedback cycles, they finished the session by completing only the short suppression training phase, the Strategy Questionnaire, and the BDI-II.

Suppression phase. The main suppression phase (think–no-think) was preceded by a few trials for practice in carrying out the instructions. In both the practice and main versions, each trial was preceded by a series of crosses for 200 ms. Cue words appeared in either red or green font, centered on a white background. Participants were instructed to avoid saying or thinking about the response words associated with cue words appearing in red, which were displayed for 4 s each. Participants were asked to read and fully comprehend the cue word and to stay focused on the cue word for the entire 4-s presentation time but to avoid thinking about the associated response word. The instructions emphasized that it was imperative that participants prevent the response word from coming to mind at all and that they should not think about the response word while looking at the cue word, even after the cue word had gone off the screen. If participants mistakenly responded to a red cue word, the computer displayed a series of large red *X*s. Participants were told to recall out loud the correct responses to cue words appearing in green font, just as they had done in the learning phase. Green cue words also were displayed for 4 s (or less if the response occurred sooner).

During the practice trials, one filler cue word appeared in red eight times and six filler cue words appeared in green (once or twice each). Immediately following the offset of green cue words, the computer displayed the correct response in blue font. Participants were told to use this feedback to improve their accuracy and speed of responding to green cue words. At the end of the practice phase, participants completed a questionnaire to ensure that they understood and had followed the instructions.

The main suppression phase differed from practice trials in one respect: The blue feedback word was displayed only when participants responded incorrectly to green cue words. The suppression phase consisted of 244 trials with two filler cues at the beginning and end of the list. The trials appeared in a randomized block order with six filler cue words each repeated 12 times (72 trials), six cue words for responding repeated 12 times, six cue words for suppressing repeated 12 times, another six "respond" cue words repeated twice (12 trials), and another six "suppress" cue words repeated twice.[1] The first presentations of the twice-presented cue words occurred in the first half of the trials. All cue words within sets presented 12 times occurred before the next random iteration began.

Final recall test. In this phase, participants were instructed to recall all response words from the learning phase, regardless of whether they had suppressed or responded to the cue words in the suppression phase. Each cue word appeared in black font centered on white background for 4 s, or less if the participant responded sooner. Trial order consisted of four filler words followed by six randomized blocks of one cue from each of the six counterbalancing sets.

Results

Overall Analysis

The percentages of response words recalled on the final test were analyzed by a mixed-design analysis of variance (ANOVA), with between-subjects factors of group (MDD and control) and the valence of suppressed–responded responses (suppress-positive–respond-negative vs. suppress-negative–respond-positive), and within-subject factors of instruction (suppress vs. respond) and the number of cue presentations (0, 2, or 12) during the suppression phase. We evaluated the linear trend across the number of cue presentations to examine the question of below-baseline suppression. *Below-baseline suppression* refers to the decreased rate of recall of response words that were associated with cues that were presented 2 or 12 times during the suppression phase compared with response words that were associated with cues that were never presented in the suppression phase. We also included coun-

[1] Anderson and Green (2001) used 1, 8, and 16 repetitions. We reduced the number of repetitions to 2 and 12 to shorten the experiment to make it more reasonable for community volunteers and for the clinical sample.

terbalancing condition as a between-subjects factor as a strategy for reducing error variance. Significant main effects and interactions that are qualified by significant higher order interactions are not reported.

The suppression effect found by Anderson and Green (2001) and subsequent replications are expressed as the interaction of the linear component of cue word presentations (baseline vs. 12 presentations, in this case) with instruction to suppress or respond during the suppression phase. That significant interaction was qualified by the three-way interaction with valence, $F(1, 48) = 4.54$, $MSE = 190.97$, $p < .04$. Participants were more successful in forgetting negative words as a function of suppression practice. (Although diagnostic group did not interact significantly with the effect, results reported in the next section show that this conclusion applies mainly to the recall by MDD participants.)

The only other nonqualified significant effect in the overall analysis was the interaction of diagnostic group and valence, $F(1, 48) = 6.15$, $MSE = 652.65$, $p < .02$. Overall, the emotional valence of the response words did not seem to matter for control participants. The mean percentage of response words recalled was 84% by those who suppressed positive and responded with negative versus 83% by those who suppressed negative and responded with positive words. In contrast, MDD participants recalled an average of 85% of the words if they suppressed positive and responded with negative words but only 73% if they suppressed negative and responded with positive words. To show how each group performed in the full design, we present the group means separately: control group means in Figure 1 and MDD group means in Figure 2.[2]

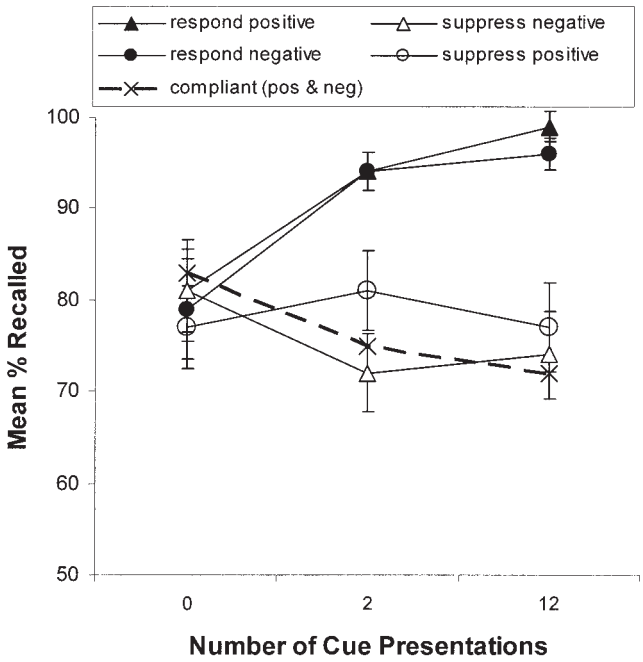

Figure 1. Mean percentage of response words recalled by the control participants as a function of instruction (suppress–respond), valence (positive–negative), and number of cue presentations. Participants were instructed to respond to positive and to suppress negative words or vice versa. Error bars = 1 *SE*. pos = positive; neg = negative.

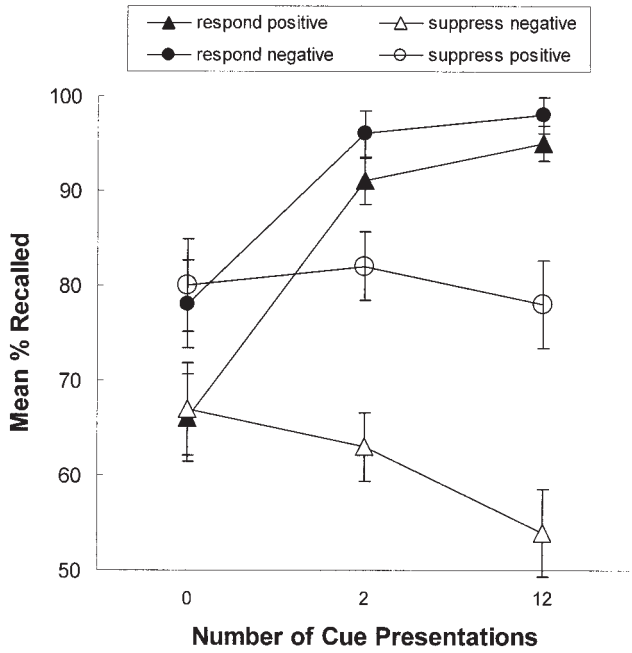

Figure 2. Mean percentage of response words recalled by the participants diagnosed with major depressive disorder (MDD) as a function of instruction (suppress–respond), valence (positive–negative), and number of cue presentations. Participants were instructed to respond to positive and to suppress negative words, or vice versa. Error bars = 1 *SE*.

Nondepressed Control Participants

Our first goal was to examine whether we would be able to replicate Anderson and Green's (2001) finding of below-baseline recall after suppression practice in our control group using our modified design. Thus, we predicted a significant interaction within the control group of instruction (to suppress or to respond) and the number of cue presentations. As shown in Figure 1, this interaction was significant, $F(1, 24) = 21.81$, $MSE = 187.11$, $p < .001$. There were, however, no significant effects within the suppressed items only, $F(1, 70) < 1$, indicating that the nondepressed controls did not exhibit below-baseline suppression. This outcome might be considered a failure to replicate conceptually the results reported by Anderson and Green if we assume that their participants were also not depressed. A fairer test of a conceptual replication of Anderson and Green's results, however, is one that excludes data from participants who reported having used strategies, indicating that they did not comply with instructions.

Accordingly, we computed a noncompliance score by summing the ratings for the first three items on the strategy questionnaire. Then, within each counterbalancing cell, we excluded the data

[2] Although our design differs in a number of ways from studies on mood congruency in depression, we found indications of mood-congruent recall when comparing depressed and control participants' recall of unpracticed words in the respond condition. Whereas the control participants recalled equal numbers of positive and negative unpracticed words, $F(1, 24) < 1$, the depressed participants show a trend toward recalling more negative than positive unpracticed words, $F(1, 24) = 3.45$, $p < .08$.

from the individual with the highest noncompliance score of the three, and reanalyzed the data from the remaining 24 control participants (6 counterbalancing conditions × 2 valence conditions). As predicted, independent of valence, significant below-baseline suppression was obtained, $F(1, 12) = 6.43$, $MSE = 202.55$, $p < .03$. These means are represented by a dashed line in Figure 1. The mean noncompliance score was 3.5 for the 12 excluded participants and 1.1 for the remaining 24 participants, $t(34) = 5.11$, $SE = 0.47$, $p < .001$. As expected, the noncompliance score was negatively correlated with the size of the suppression effect (percentage recalled from responded items minus percent recalled from suppressed items), $r(34) = –.34$, $p < .05$.

MDD Participants

The primary goal of this study was to examine whether depressed participants who are trained to suppress negative target words would exhibit below-baseline recall of these words. Thus, we expected to obtain an interaction of instruction, cue presentation, and valence. As is apparent in Figure 2, the three-way interaction of instruction, valence, and the linear component of cue word presentations (baseline vs. 12 presentations) was significant for the MDD participants, $F(1, 24) = 4.79$, $MSE = 194.83$, $p < .04$.[3] Below-baseline recall was clearly obtained by those depressed participants who practiced suppressing negative words, $F(1, 12) = 8.17$, $MSE = 185.18$, $p < .025$, but not by those depressed participants who practiced suppressing positive words, $F(1, 12) < 1$. In addition, these participants' baseline recall, almost identical in the two baseline conditions (suppress vs. respond 0 times), was significantly lower than the baseline recall of MDD participants who suppressed positive words. In summary, the MDD participants were not only more successful in forgetting negative words that they had practiced suppressing, but they were also less successful in remembering all unpracticed (baseline) words and equally successful in recalling practiced responses, regardless of valence.

It is interesting that the MDD participants who practiced suppressing negative words also recalled fewer of them than did the comparable control participants (*M*s = 61% vs. 76%, respectively); the two diagnostic groups recalled similar levels of positive words that they had suppressed (MDD: $M = 80\%$; controls: $M = 78\%$). Indeed, the interaction of group and valence within suppressed responses was significant, $F(1, 48) = 5.41$, $MSE = 643.00$, $p < .03$.

Overall, there were no significant differences between depressed and nondepressed participants in the suppress-positive–respond-negative condition. However, depressed and nondepressed participants in the suppress-negative–respond-positive condition differentially recalled baseline (unpracticed) words, $F(1, 24) = 5.93$, $MSE = 625.00$, $p < .03$ (*M*s = 81% by nondepressed vs. 66% by depressed participants). Last, compared with their nondepressed counterparts, these depressed participants also recalled many fewer negative responses associated with cues presented 12 times during practice (*M*s = 54% vs. 74%, respectively), $F(1, 24) = 12.10$, $MSE = 308.64$, $p < .002$.

Group Difference in Strategy

Noncompliance. An ANOVA conducted on the noncompliance scores with factors for group and valence yielded a significant main effect for valence, $F(1, 68) = 6.15$, $MSE = 3.44$, $p < .02$ (*M*s= 1.4 in the suppress-negative condition and 2.5 in the suppress-positive condition). Neither the main effect for group nor the interaction of valence and group was statistically significant (both *F*s < 1).

Substitution and mind-blanking strategies. Similar analyses were conducted separately on the ratings for Questions 4 (thinking about a substitute for the original response) and 5 (mind blanking) of the Strategies Questionnaire. Although the participants reported that they "rarely" to "sometimes" used mind blanking ($M = 1.5$), they reported frequent use of thought substitution to keep from thinking about the response word ($M = 3.0$). No significant effects were obtained for either item.

Discussion

The primary purpose of the present study was to examine whether depressed participants can actively and intentionally forget negative material. We explicitly instructed depressed and nondepressed participants to not think about emotional words that they had previously learned to associate with neutral cue words, and we gave them multiple occasions to practice suppression. We found that depressed participants were particularly successful in forgetting negative responses. In fact, the depressed participants recalled negative words that they had suppressed significantly less often than they recalled negative words in the baseline condition. Furthermore, the more that the depressed participants practiced suppressing the negative items, the fewer items they recalled on the final test. After controlling for noncompliance with the suppression instructions, we also found evidence for below-baseline recall by control participants but to a comparable degree for positive and negative responses. Thus, this is the first study to replicate the results reported by Anderson and Green (2001) using emotional material.

This study is also one of the first to investigate intentional forgetting of emotional material in depression. So far, in studies investigating memory in depression, forgetting is considered a passive process that reflects individual differences in the elaboration of the presented material. Thus, in the majority of these studies, participants are not explicitly instructed to forget emotional material. In studies examining passive forgetting, depression is typically associated with better memory for emotionally negative material, presumably because depressed individuals habitually attend to and elaborately process negative information. Indeed, although our design differs in a number of important ways from mood-congruency paradigms (e.g., learning to criterion, practicing recall or suppression), control participants recalled equal numbers of positive and negative unpracticed words in the "respond" condition, whereas the depressed participants showed a trend toward recalling more negative than positive unpracticed words. Despite such habits, the clinically depressed participants in the present study were successful in forgetting negatively valenced words. In contrast to other studies on memory in depression, we explicitly

[3] A similar difference in compliance means was obtained in the group of MDD participants. The corresponding reanalysis of the data in the suppressed condition revealed below-baseline suppression only for those participants who practiced suppressing negative responses, $F(1, 6) = 7.14$, $MSE = 162.04$, $p < .05$.

instructed our participants to stop the retrieval of emotional nouns and gave them the opportunity to practice suppressing the negative response words. Depressed participants recalled fewer negative words than did control participants and fewer than did other depressed participants who suppressed positive words.

This pattern of results stands in contrast to earlier findings in studies of directed forgetting. For example, in the only study to date that has examined directed forgetting for emotional words in depression, Power et al. (2000) reported that depressed participants recalled more to-be-forgotten negative than positive words and showed better recall for to-be-forgotten than for to-be-remembered negative words. However, Power et al. found this effect only in one of their three studies, in which participants were asked to make self-reference judgments about the words. One significant difference between our study and Power et al.'s investigation, therefore, is that we did not have participants encode the stimuli self-referentially. Equally important, the degree to which the directed forgetting paradigm represents intentional forgetting is questionable. Whereas the instruction to forget is given only once in directed forgetting studies, the think–no-think design provides multiple occasions for practicing suppression in response to cues previously used to remember. Indeed, the effect of the number of cue presentations in our results demonstrates that practice is an important aspect of intentional forgetting.

Our results also differ from those reported by Hertel and Gerstle (2003), who used a similar think–no-think design. In that study, dysphoric students recalled more words that they had been instructed to forget than did control participants and failed to produce below-baseline forgetting in response to both positive and negative cues. Hertel and Gerstle's method differed from ours in so many potentially important ways that it is difficult to speculate about the reasons for the differences in outcomes. Nevertheless, some of the following differences should be considered in future research: (a) Hertel and Gerstle's participants were dysphoric students, whereas participants in the present study were older and were carefully diagnosed with MDD. (b) Hertel and Gerstle did not use a Strategy Questionnaire, so noncompliance with instructions for suppression cannot be ruled out as occurring in the dysphoric group, leading to better memory performance. (c) Hertel and Gerstle's learning phase involved the self-referential imaging of items made positive or negative by the adjective. Arguably, forgetting these self-generated, integrated, and elaborated episodes might be harder than forgetting the imposed experimental associations between the unrelated noun pairs used in the present study. (d) The cue words used in Hertel and Gerstle's study were emotional in their own right (and the responses to be suppressed were at least nominally neutral), so any emotional quality of the memory to be suppressed was always present on each suppression trial. In the present study, the cue words were neutral and the emotional quality of the memory was carried by the response to be suppressed, which might have permitted valence differences to emerge.

Considerable research on intentional thought suppression has been guided by the notion of ironic control processes (Wegner, 1994), whereby attempting to suppress or push away an unwanted thought can increase its accessibility, particularly under high cognitive load. Although the results of some of the studies in this area suggest that thought suppression is counterproductive (Beevers & Meyer, 2004; Wegner, Erber, & Zanakos, 1993; Wenzlaff & Wegner, 2000), investigators explicitly examining recall have found that suppression leads to poorer recall of the suppressed stimuli (Rassin, 2001; Wegner, Quillian, & Houston, 1996), a finding consistent with the present data. For the control participants in our study the finding of below-baseline recall in the suppression condition was obtained only after taking into account participants' differential compliance with the suppression instructions. This is consistent with Anderson's (personal communication, May 18, 2002) findings (see also Hertel & Calcaterra, 2005) and supports the notion that suppression training is indeed necessary for forgetting. In addition, it also suggests that the outcome of suppression is difficult to achieve. Moreover, our current design does not allow us to exclude the possibility that the forgetting observed after suppression training is due primarily to interference from substitute words or pictures that participants used to prevent thinking about the response word. Indeed, Hertel and Calcaterra recently provided evidence that the use of a self-initiated strategy of thought substitution can increase the degree of forgetting in the think–no-think task. Nevertheless, the finding that depressed participants can intentionally forget negative material is important and novel, regardless of whether it is due to the active inhibition of memory representations or to the interference from substitute thoughts or images. Future studies might successfully address the specific mechanisms that underlie the increased forgetting of negative words in the depressed group.

Two relatively trivial explanations of the present results should also be considered here. First, the finding that depressed participants showed no deficit in intentional forgetting is equivalent to the finding that depressed participants show deficient recall in general. Therefore, it is important to consider whether the results were due to a general depression-related deficit in recall (Burt, Zembar, & Niederehe, 1995). Overall, more nondepressed than depressed participants were excluded because they did not reach the criterion in the learning phase. The depressed and nondepressed participants who remained in our sample did not differ in the number of learning trials required for reaching the criterion ($M = 2.1$ vs. $M = 1.9$, respectively; $t < 1.0$). In addition, the depressed and control participants in both valence conditions had equivalent rates of recall following practice in responding during the think–no-think phase (although ceiling effects should be suspected after 12 practice opportunities). Most important, however, are the similar levels of recalling the positive words that the two groups practiced suppressing. Evidence that deficient recall in depression was restricted to suppressed negative responses is not compatible with the notion of a general deficit.

A second explanation involves the materials used in this study. Whenever items to be recalled differ between valence conditions, the outcome concerning valence is correlational; we cannot be sure that differences other than valence are not responsible for differing results. For example, it is possible in this study that intrapair associations differed in some unanticipated way. However, any viable explanation for the present results that involves materials could not be straightforward, because significant valence differences were not found in any condition in the control group or in depressed participants' recall in the baseline and respond conditions. This pattern of results cannot be explained by a simple "materials" account. Nevertheless, our conclusions are restricted to the use of emotional nouns initially learned under conditions that did not require self-reference. Although mood-congruency effects

have been found with a variety of tasks, involving both self-referential and semantic processing, stronger effects have been reported when depressed participants are required to process the words self-referentially and when adjectives, rather than nouns, are presented (for an overview see Matt et al., 1992). Thus, we cannot rule out the possibility that suppressing words that have been processed with reference to the self is more difficult for the depressed than for the nondepressed participants. For this first study of practiced suppression of emotional words by a clinical sample, however, it was important to replicate many features of the method used by Anderson and Green (2001). In that regard, it was important to ensure that the noun pairs were not related a priori, and it would be difficult to ask participants to process such pairs self-referentially. Questions concerning self-referential processing and memory for adjectives should be addressed in future research.

It is important to note that our results suggest that intentional forgetting of negative material in depression comes at a cost of forgetting baseline words, for which the cues were not presented during the think–no-think phase. Compared with the controls and with the depressed participants who practiced suppressing positive words, depressed participants who practiced suppressing negative words showed significantly poorer recall of the unpracticed baseline words. This finding might be related to the strategies used by participants to achieve suppression. Hertel and Calcaterra (2005) found that intentional forgetting benefits substantially from thought substitution during the suppression phase. Depressed people are particularly prone to rumination, which consists largely of relational thinking about negative events. And there is some possibility that depressed persons tend to distract themselves from unwanted negative thoughts using other negative thoughts (Wenzlaff, Wegner, & Roper, 1988). Therefore, the suppression of negative words might be differentially enhanced by thought substitution in depression. We did not find significant differences in the reported frequency of thought substitution; instead, participants in all conditions reported frequent use of this strategy. However, although we assessed whether the participants used thought substitution in the suppression phase, we did not assess which alternative thoughts they generated and whether they used positive or negative words to distract themselves from the response word. This may be an important addition for future research. Thought substitution merits further exploration in future studies of intentional forgetting, perhaps with methods other than self-report. Self-reports are often insensitive measures of cognitive phenomena and should be similarly insensitive to the extent of actual substitution during the earlier suppression phase. In using other methods, if a greater tendency to substitute other thoughts emerged in the condition in which the depressed participants' task was to suppress negative responses, that tendency might impair recall of the unpracticed words in a manner analogous to retrieval-induced forgetting. In retrieval-induced forgetting, practicing the retrieval of material that competes with unpracticed memories leads to reduced recall of the unpracticed memories (Anderson & Spellman, 1995). Of course, this hypothesis is highly speculative, because the basis on which the thought substitutes would compete with the studied items is unclear.

Still, this leaves us with an intriguing question: If depressed individuals are able to intentionally forget negative material, why is recurring negative thoughts and negative memories such a prevalent symptom of this debilitating disorder? Habits of thought might cause negative thoughts to come to mind, but continued attention and awareness might be under more intentional control than we have previously realized (Hertel, 2004). The tendency of depressed individuals to ruminate about negative information instead of engaging in intentional forgetting may play an important role in explaining why mood-congruent recall is observed so reliably in depression. Consistent with this suggestion, Watkins (2002) argued that rumination is a form of conceptual elaboration of unpleasant thoughts that leads eventually to enhanced recall, not only of the encoded material but of related material as well. Similarly, McFarland and Buehler (1998) found evidence of mood-congruent recall when participants ruminated in response to a negative mood induction, but not when they engaged in self-reflective thinking. It is interesting that Papageorgiou and Wells (2003) recently emphasized that rumination might be more controllable than previously assumed and that depressed individuals' metacognitive beliefs about rumination might play an important role in maintaining their ruminative tendencies. More specifically, Papageorgiou and Wells suggested that at the same time that depressed individuals hold positive beliefs about rumination as a coping strategy, they hold negative beliefs about the uncontrollability of rumination. Thus, depressed individuals might deliberately engage in rumination in an attempt to solve their problems but then become overwhelmed by negative thoughts about their ruminations. The depressive tendency to ruminate instead of to engage deliberate thought suppression or distraction when confronted with negative events certainly deserves further investigation (see Lyubomirsky & Nolen-Hoeksema, 1993), as does the possibility that this tendency can be overcome by training (see Teasdale et al., 2000). The present results suggest that training depressed individuals in intentional forgetting could prove to be an effective strategy to counteract their automatic ruminative tendencies.

References

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

Anderson, M. C., & Green, C. (2001). Suppressing unwanted memories by executive control. *Nature, 410,* 366–369.

Anderson, M. C., & Spellman, B. A. (1995). On the status of inhibitory mechanisms in cognition: Memory retrieval as a model case. *Psychological Review, 102,* 68–100.

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

Beevers, C. G., & Meyer, B. (2004). Thought suppression and depression risk. *Cognition and Emotion, 18,* 859–867.

Bradley, M. M., & Lang, P. J. (1999). *Affective norms for English words (ANEW): Technical manual and affective ratings.* Gainesville, FL: Center for Research in Psychophysiology, University of Florida.

Burt, D. B., Zembar, M. J., & Niederehe, G. (1995). Depression and memory impairment: A meta-analysis of the association, its pattern, and specificity. *Psychological Bulletin, 117,* 285–305.

Charles, S. T., Mather, M., & Carstensen, L. L. (2003). Aging and emotional memory: The forgettable nature of negative images for older adults. *Journal of Experimental Psychology: General, 132,* 310–324.

First, M. B., Spitzer, R. L., Gibbon, M., & Williams, J. B. W. (1996). *Structured Clinical Interview for DSM–IV Axis I Disorders-Clinician Version (SCID-CV).* Washington, DC: American Psychiatric Press.

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.

Kennedy, Q., Mather, M., & Carstensen, L. L. (2004). The role of motivation in the age-related positivity effect in autobiographical memory. *Psychological Science, 15,* 208–214.

Korfine, L., & Hooley, J. M. (2000). Directed forgetting of emotional stimuli in borderline personality disorder. *Journal of Abnormal Psychology, 109,* 214–221.

Levy, B. J., & Anderson, M. C. (2002). Inhibitory processes and the control of memory retrieval. *Trends in Cognitive Sciences, 6,* 299–305.

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.

Lyubomirsky, S., & Nolen-Hoeksema, S. (1993). Self-perpetuating properties of dysphoric rumination. *Journal of Personality and Social Psychology, 65,* 339–349.

Matt, G. E., Vazquez, C., & Campbell, W. K. (1992). Mood-congruent recall of affectively toned stimuli: A meta-analytic review. *Clinical Psychology Review, 12,* 227–255.

McFarland, C., & Buehler, R. (1998). The impact of negative affect on autobiographical memory: The role of self-focused attention to moods. *Journal of Personality and Social Psychology, 75,* 1424–1440.

Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (1998). The University of South Florida word association, rhyme, and word fragment norms. Retrieved May 30, 2003 from http://www.usf.edu/FreeAssociation/

Nolen-Hoeksema, S. (2000). The role of rumination in depressive disorders and mixed anxiety/depressive symptoms. *Journal of Abnormal Psychology, 109,* 504–511.

Nolen-Hoeksema, S., & Larson, J. (1999). *Coping with loss.* Hillsdale, NJ: Erlbaum.

Nolen-Hoeksema, S., Morrow, J., & Fredrickson, B. L. (1993). Response styles and the duration of episodes of depressed mood. *Journal of Abnormal Psychology, 102,* 20–28.

Papageorgiou, C., & Wells, A. (2003). An empirical test of a clinical metacognitive model of rumination and depression. *Cognitive Therapy & Research, 27,* 261–273.

Power, M. J., Dalgleish, T., Claudio, V., Tata, P., & Kentish, J. (2000). The directed forgetting task: Application to emotionally valent material. *Journal of Affective Disorders, 57,* 147–157.

Rassin, E. (2001). Thought suppression, memory, and interrogative suggestability. *Psychology, Crime & Law, 7,* 45–55.

Roberts, J. E., Gilboa, E., & Gotlib, I. H. (1998). Ruminative response style and vulnerability to episodes of dysphoria: Gender, neuroticism, and episode duration. *Cognitive Therapy and Research, 22,* 401–423.

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

Teasdale, J. D., Segal, Z. V., Williams, J. M. G., Ridgeway, V. A., Soulsby, J. M., & Lau, M. A. (2000). Prevention of relapse/recurrence in major depression by mindfulness-based cognitive therapy. *Journal of Consulting and Clinical Psychology, 68,* 615–623.

Tolin, D. F., Hamlin, C., & Foa, E. B. (2002). Directed forgetting in obsessive-compulsive disorder: Replication and extension. *Behaviour Research & Therapy, 40,* 792–803.

Watkins, P. C. (2002). Implicit memory bias in depression. *Cognition and Emotion, 16,* 381–402.

Wegner, D. M. (1994). Ironic processes of mental control. *Psychological Review, 101,* 34–52.

Wegner, D. M., Erber, R., & Zanakos, S. (1993). Ironic processes in the mental control of mood and mood-related thought. *Journal of Personality and Social Psychology, 65,* 1093–1104.

Wegner, D. M., Quillian, F., & Houston, C. F. (1996). Memories out of order: Thought suppression and the disturbance of sequence memory. *Journal of Personality and Social Psychology, 71,* 680–691.

Wenzlaff, R. M., & Wegner, D. M. (2000). Thought suppression. *Annual Review of Psychology, 51,* 59–91.

Wenzlaff, R. M., Wegner, D. M., & Roper, D. W. (1988). Depression and mental control: The resurgence of unwanted negative thoughts. *Journal of Personality and Social Psychology, 55,* 882–892.

Williams, J. M. G., Watts, F. N., MacLeod, C., & Mathews, A. (1997). *Cognitive psychology and emotional disorders* (2nd ed.). New York: Wiley.

Received September 2, 2004
Revision received April 29, 2005
Accepted May 7, 2005 ■