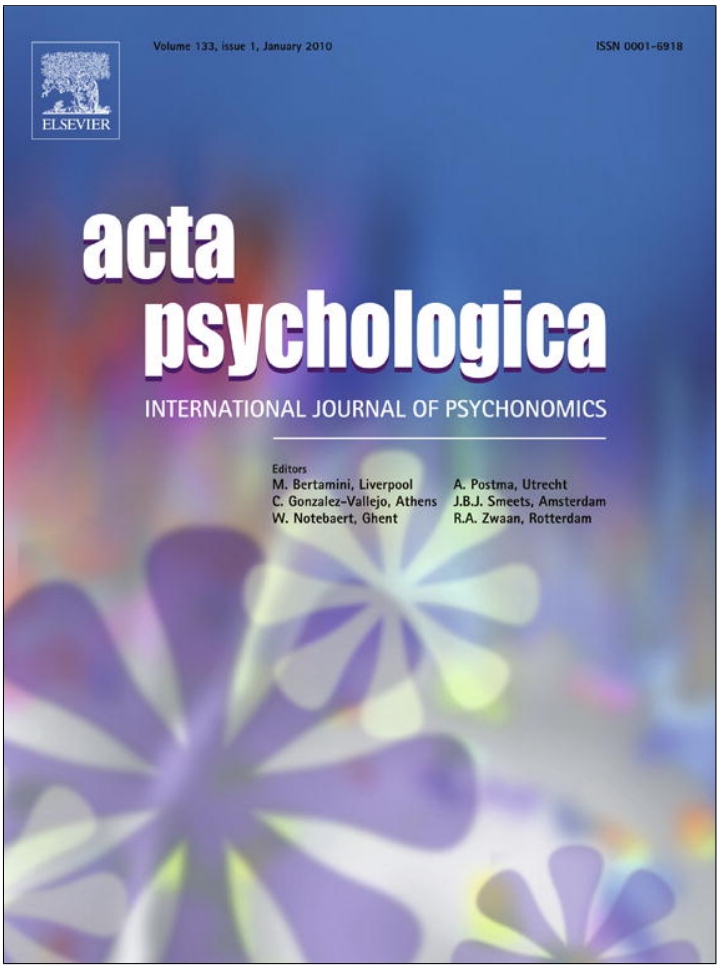

Acta Psychologica 133 (2010) 73–80

Contents lists available at ScienceDirect

Acta Psychologica

journal homepage: www.elsevier.com/locate/actpsy

Forgetting to forget: On the duration of voluntary suppression of neutral and emotional memories

Simon Nørby *, Martin Lange, Axel Larsen

Department of Psychology, University of Copenhagen, Denmark

ARTICLE INFO

Article history:
Received 6 November 2008
Received in revised form 30 September 2009
Accepted 2 October 2009
Available online 10 November 2009

PsyclNFO classification:
2343
2360

Keywords:
Think/no-think
Forgetting
Delayed recall
Emotion

ABSTRACT

Can we control the content of our memory and forget what we do not want to think about by an act of will? If so, is forgetting temporary or permanent, and is it independent of the nature of what we wish to forget? Using Anderson and Green's (2001) "think/no-think" paradigm with neutral and emotional nouns, we found in agreement with other studies that memory for neutral words was reduced instantly upon repeated attempts at suppression. However, the effect was temporary and vanished after a period of one week, which strongly suggests that intended memory suppression interferes with immediate retrieval but does not lead to long-term forgetting. Furthermore, the amount of training that clearly reduced immediate recall of neutral items did not at all reduce recall of emotional items. This finding is in accordance with the notion that emotional items have a higher degree of salience and tend to attract more attention than neutral items.

1. Introduction

Can we forget unwanted memories by an act of will? Recent research suggests that it is possible to disregard recently encoded memories by the use of memory control strategies (see Bjork, Bjork, & MacLeod, 2006; Levy & Anderson, 2002). However, the generality of these findings remains to be settled and at least two issues ought to be addressed before the question about whether volitional forgetting can take place on unwanted memories may be properly answered. First, it should be determined if memory control strategies have more than a temporary effect on to-be-forgotten memories. If they do not, such strategies would appear to be rather futile in daily life where people may wish to forget unwanted memories over an extended period of time, or even permanently. Second, it should be examined if memory control strategies have a reliable effect on negative memories. If this is not the case, they may be rather ineffective on what are perhaps the most unwanted memories of all, that is, unpleasant recollections of past failures and troubling incidents. All recent experiments on volitional forgetting have to our knowledge investigated short-term effects and most have employed neutral stimuli material. Therefore, the primary objective of the present study was to investigate the long-term effect of volitional forgetting by employing a delayed re-test. A secondary objective was to further examine our ability to forget negative memories on the basis of some contradictory results concerning valence and forgetting (e.g., Depue, Banich, & Curran, 2006; Marx, Marshall, & Castro, 2008).

Several recent experimental paradigms suggest that we may be capable of forgetting recently encoded memories on purpose. In the popular think/no-think (TNT) paradigm, which we employed in the present study, participants are instructed to actively avoid that certain memories enter their minds. First, they learn a number of cue-target pairs. Then, in the critical phase, they are continuously exposed to cues and instructed to suppress the associated targets in relation to some of these cues (no-think condition) while rehearsing the targets for the rest of them (think condition). When tested later on, recall of items attempted suppressed is typically inferior to recall of a group of studied baseline targets not involved in the critical phase of the experiment (e.g., Anderson & Green, 2001; Anderson et al., 2004; Depue, Curran, & Banich, 2007; Depue et al., 2006; Hertel & Calcaterra, 2005; Levy & Anderson, 2008; however, see Bulevich, Roediger, Balota, & Butler, 2006). In the related list-method directed forgetting (DF) paradigm, the participants are instructed to ignore a previously studied list of items

* Corresponding author. Address: Department of Psychology, University of Copenhagen, Øster Farimagsgade 2A, 1353 Copenhagen K., Denmark. Tel.: +45 26236712.
E-mail address: simon.noerby@psy.ku.dk (S. Nørby).

0001-6918/$ - see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.actpsy.2009.10.002

(list one) and focus on learning a new list. Experiments with this paradigm usually show that recall of list one items is inferior to recall of list one items among the participants in a control group who have been told to remember both lists of words (for reviews, see Bjork et al., 2006; MacLeod, 1998).

Suppression of memories that are not relevant to current action goals may be a prerequisite for optimizing behaviour for at least two reasons. First, it may help update the memory system when a learned response is replaced because it is invalid and no longer adequate (Bjork, 1989). For example, when trying to remember a new pin code for a credit card we cannot act quickly and adequately if we cannot avoid interference by effectively suppressing the memory of the old pin code. Second, it may sustain goal-directed information processing when a recollection is stopped altogether because it is distracting or unwanted (Anderson & Green, 2001; Levy & Anderson, 2002). For example, the prospect of attending a job interview might bring to mind memories of failed job applications and one may want to suppress such recollections so that it is possible to prepare properly for the interview. Results from the use of TNT and list-method DF paradigms where performance on memory tests fall below baseline conditions indicate that memory control strategies may succeed and that suppression may decrease recall of invalid or unwanted memories.

At this point it is not known whether observed decrements in recall are stable or transient despite the possibility that the answer to this question could have important clinical implications (see Gleaves, Smith, Butler, & Spiegel, 2004). However, some observations suggest that such decrements could be transitory. First, thought suppression requires mental effort, so when concentration slips or suppression is no longer reinforced, there is reason to believe that suppressed thoughts can re-emerge (see e.g., Wegner, 2009; Wenzlaff & Wegner, 2000). Second, the ubiquitous phenomenon of spontaneous recovery shows that material that cannot be recalled on an immediate test may be recalled on a delayed test. Spontaneous recovery has been demonstrated in a number of studies in which subjects have learned two paired-associate A–B, A–C lists and exhibited recovery of B items on delayed tests (for reviews, see Brown, 1976; Wheeler, 1995). As shown by others (Hertel & Calcaterra, 2005), the TNT paradigm may bear some resemblance to classical A–B, A–C paradigms.[1]

The principal goal of the present study was thus to investigate the duration of intended memory suppression in the TNT paradigm. A secondary objective was to explore the efficiency and duration of memory control of emotional items, and in particular to check a simple observation that seems to explain why some experiments on suppression of emotional memories report that such memories may be suppressed (Depue et al., 2006, 2007; Joormann, Hertel, Brozovich, & Gotlib, 2005), while other experiments report that such memories are difficult to suppress (Marx et al., 2008; see also Hertel & Gerstle, 2003). Specifically, it seems that memory suppression is found whenever participants with a high degree of confidence can predict the occurrence of emotional items, but not when the occurrence of emotional items is random and cannot be prepared for.

For historical, theoretical as well as practical reasons forgetting of emotional and especially negative material has particular interest. Negative material may be said to be unwanted in a strong sense because it is related to matters of personal concern. This is not the case with neutral material that is at the most undesired because it might distract current information processing, i.e., unwanted in a weak sense. One view suggests that we should be able to suppress negative emotional memories that are unpleasant and inconvenient, and that voluntary forgetting may succeed on emotional material (Conway, 2001; Depue et al., 2006; Levy & Anderson, 2008). In accordance with this view, Depue et al. (2006) employed the TNT paradigm with face-word pairs (Experiment 1) and face-picture pairs (Experiment 2) and found the largest suppression effect to be present on negative relative to neutral items. Also, Depue et al. (2007) employed the TNT paradigm in a fMRI study with face-picture pairs and found that negative pictures could be suppressed (no neutral control condition was employed). Finally, Joormann et al. (2005) reported evidence that positive as well as negative words were suppressed to a similar degree in the TNT paradigm. In sum, these studies suggest that negative memories may be suppressed, at least in some circumstances.

Another view suggests quite the contrary, namely, that it should be difficult to suppress negative emotional memories. Emotional stimuli presumably capture attention more easily and lead to more intensive processing than neutral stimuli (for a review, see Compton, 2003). Also, dedicated neurobiological processes seem to support the formation and storage of emotional memories (see LaBar & Cabeza, 2006). Finally, emotional experiences may be rehearsed more often than neutral experiences because the former but not the latter are personally significant. Therefore, one could expect emotional material to be encoded and consolidated better than neutral material. In accordance with this, many studies have found that emotional memories are retained better than neutral memories (for reviews, see Levine & Pizarro, 2004; Reisberg & Heuer, 2004). One reason may be that we have an evolutionary-based preference for emotional information. Specifically, negative emotions such as fear and disgust should support survival because they signal objects and situations that generally pose a threat and should be avoided. This suggests that negative emotional material should be fairly resistant to voluntary forgetting. Indeed, one recent study with the TNT paradigm found that negative target words were not suppressed (Marx et al., 2008; see also Hertel & Gerstle, 2003).

We conjecture that Depue et al.'s (2006) results were due to a particular feature in their study, namely, the use of a blocked design in which neutral and emotional stimuli were presented in separate blocks of trials, which may have caused habituation by making emotional items entirely predictable. Also, the results of Depue et al. (2007) and Joormann et al. (2005) may be explained in a similar fashion (see below for an elaboration on this point). Habituation consists of a decrease in responsiveness to repeated presentations of the same stimulus or stimuli of the same type. Subjects have been found to habituate to continuous presentations of emotional stimuli as measured by, for example, decreased activity in critical brain regions (e.g., Feinstein, Goldin, Stein, Brown, & Paulus, 2002; Wright et al., 2001) and of psychophysiological responses (e.g., Bradley, Lang, & Cuthbert, 1993). In the experiment reported here, we attempted to avoid habituation by mixing neutral and emotional items.

In sum, to test the generality of volitional forgetting with respect to duration and type of material, the specific goals of the present study were to determine whether intended memory suppression of neutral material is stable – and still measurable when a week has passed – and whether intended memory suppression of negative material can be counteracted or at least reliably reduced compared to neutral material by simply making the occurrence of negative stimuli unpredictable. We used the TNT paradigm and tested highly practiced suppression of neutral and negative nouns, both immediately following practice and after a retention interval of one week.

[1] A–B, A–C paradigms have typically been used to study retroactive interference in immediate tests. One recent study demonstrated that short-term retroactive interference may also be produced in the TNT paradigm (Hertel & Calcaterra, 2005). This study provided participants with a substitute target to use throughout the TNT phase and thereby created a sort of A–B, A–C paradigm. However, the present discussion should be understood in the context of studies that have employed delayed tests and found spontaneous recovery. Such results have been interpreted in terms of a decreasing influence of suppression on B items over time (e.g., Bjork, 2001; Postman, Start, & Fraser, 1968).

2. Method

2.1. Participants

Fifty-nine healthy Danish-speaking students or recent graduates (49 females) between 18 and 30 years (mean age = 24.4) were recruited through posters and email correspondence. The participants were informed that the experiment was about memory but were naïve as to the specific purpose of the experiment. They were asked to attend on two occasions separated by a one week interval (plus or minus four hours), producing a total testing time of between three and five hours. The participants were paid by the hour.

Six participants were excluded because they did not fulfill the initial learning criteria of 60% (see below) and five participants were excluded because they did not comply with the instructions, leaving a final sample of 48 persons that participated in the TNT task (mean age = 24.0). In agreement with other studies (Bulevich et al., 2006; Joormann et al., 2005), exclusion by non-compliance was based on a questionnaire devised by Anderson and colleagues for this purpose (Anderson, personal communication, June, 2006). In this questionnaire which was a translation of one used by Anderson and colleagues, the subjects were asked if they – contrary to the instructions – tried to remember those target words that they were supposed to suppress. On a scale from 0 (never) to 4 (very often), they were asked to rate if they: (1) quickly checked to see if they remembered the suppress word when exposed to the cue word; (2) checked to see if they still remembered the suppress word after the cue word disappeared from the screen; or (3) thought about the suppress word that went with the cue word to improve the memory for that word pair when seeing the cue word. The subjects that had a total non-compliance score of 6 or more, or a score of 4 regarding one of the three questions, were excluded.

A sub-sample consisting of the last 30 of the 48 persons (26 females; mean age = 24.1) was subjected to a re-test after an interval of one week. When the project was initiated a re-test was employed after one month. However, this procedure resulted in floor effects. Therefore, the procedure was changed so that all the remaining participants were re-tested after one week.

2.2. Materials

The stimulus set consisted of Danish nouns with 4 to 10 letters (see Appendix). One hundred and forty words with moderate frequency were selected from the Danish "Korpus 2000" database (see Andersen, Asmussen, & Asmussen, 2002). Care was taken to ensure that the mean frequency of words selected for use in the different experimental conditions was approximately the same and that the words were neither very unusual nor very common. To avoid any effect of pre-experimental semantic association between the words, the stimulus set was constructed so that there were no synonyms or strongly related words.

Twenty-five students who did not participate in the experiment rated the words on a discrete nine-point scale ranging from −4 (*extremely negative*) to + 4 (*extremely positive*). On average, the emotional words ($-3.8 < Ms < -1.4$) were rated just about three points lower ($M = -2.59$) than the neutral words ($-0.5 < Ms < 1.5$) which were rated as marginally positive ($M = 0.49$). The emotional words had strong negative connotations concerning sickness, cruelty, obscenity, and so forth. For example, the words "leukaemia", "torture" and "incest" were used. The cue word in a word pair was always a neutral word. Half the targets were neutral and half were negative.

The design comprised a $2 \times 2 \times 2$ within-subjects design with the factors trial repetition in the critical TNT phase (8, 16), TNT instruction (suppress, respond), and valence (neutral, negative), and 2 baseline conditions (neutral, negative), the latter consisting of words not participating in the TNT phase. We used 70 word pairs. There were 5 word pairs in each of the 4 suppress and the 2 baseline conditions and 10 word pairs in each of the 4 respond conditions. Cues from 60 word pairs were presented in the TNT phase. Half were shown on 8 occasions (20 cues from to-be-responded pairs and 10 cues from to-be-suppressed pairs) and the other half were shown on 16 occasions (20 cues from to-be-responded pairs and 10 cues from to-be-suppressed pairs). No cues from the remaining 10 word pairs were shown in the TNT phase and these constituted baseline items. Also, neutral and negative word pairs were distributed evenly among conditions.

2.3. Procedure

The experiment was run with E-Prime software (Psychology Software Testing, Pittsburgh, PA) on a personal computer, and stimuli were displayed on a 17-in. (43.2-cm) screen. During the experiment (with the exception of the study phase), the experimenter sat behind the participant. The experimenter classified the participant's vocal responses by mouse input to E-Prime which then provided on-screen feedback to the participant. Galvanic skin responses were recorded in the initial learning phase. These data were uninformative and will not be reported here.

The experiment comprised four major phases in which materials were randomly assigned to experimental conditions. These phases were fairly similar to the ones employed by Anderson and Green (2001). In the first (study) phase, each participant was exposed twice to all 70 word pairs, divided into two pseudo-randomized rounds. Each word pair was presented for eight seconds in the middle of the screen, with the cue word to the left and the target word to the right. Each pair was preceded by a fixation cross shown for 250 ms and succeeded by a 3-s delay. In the second (training) phase, the individual was shown each of the cue words which were displayed on the monitor one by one for two seconds (again preceded by a fixation cross for 250 ms) and asked to orally recall the associated targets without a time limit. Feedback as to whether the participant was correct was provided and, in the case of "no" or a wrong answer, the correct target was shown. Training cycles continued until the individual reached a performance of at least 60%. If this did not happen within seven cycles, the participant was excluded from the experiment. If the learning criterion was met, a Danish translation of the Anderson and Green instructions was provided. These instructions emphasized that the participant should avoid thinking about the to-be-suppressed targets whenever he or she was exposed to the associated cues. The participant was also told about the feedback in the TNT phase (see below).

Cues from 60 randomly assigned word pairs were presented in the third (TNT) phase while cues from the remaining 10 baseline word pairs were not shown in this phase. As in the original experiment (Anderson & Green, 2001), two thirds of the trials in the TNT phase required that the participant responded with the target word by saying it out loud, and one third of the trials required that the participant suppressed and withheld the response. Furthermore, cues presented in the TNT phase were randomly selected to be presented on 8 or 16 occasions.

In the TNT phase, the participant initially learned to recognize at least 90% of the 60 cues assigned to this phase as belonging to either think or no-think pairs. This happened in up to four training cycles (in which no target words were shown). Each cycle began with a simultaneous display of the 20 cues from the randomly assigned to-be-suppressed pairs. Then, all 60 cues assigned to the TNT phase were shown one by one for 2 s (preceded by a fixation cross for 250 ms). The participant had to say whether or not each item was a cue for responding or suppressing, and was provided

with a feedback as to whether he or she was correct. In the important part of the TNT phase, the individual had to either respond or not respond to cue words displayed one by one on the screen. Each trial consisted of a fixation cross shown for 250 ms, a cue word shown for 4 s (on a suppress or a respond trial) or less (if the participant replied faster on a respond trial), and a 400-ms intertrial interval. Any response to a respond cue, as well as no response to a suppression cue, resulted in "correct" feedback. In the case of no response to a respond cue, the correct target was provided. Any response to a suppression cue resulted in "incorrect" feedback in the form of a loud error sound. Finally, in the fourth (test) phase, the individual was asked to orally report the target to each of the 70 cues irrespective of prior instructions. The cues were shown one by one for 2 s in pseudo-randomized order, preceded by a fixation cross of 250 ms. Participants had up to 30 s to respond. There was no feedback and intertrial interval was 5 s.

After completing the test phase, the participant was asked to return after one week in order to recalibrate instruments. The real purpose, to probe memory after one week, was not revealed. The re-test was similar to the test phase described above.

The present setup differed from the one used by Anderson and Green (2001) on some features. First, we used 70 word pairs and presented each word pair in two rounds for eight seconds in the study phase. In contrast, Anderson and Green used 50 word pairs (40 critical and 10 filler word pairs) and only showed each word pair once for five seconds. We used longer exposure duration and showed the word pairs twice because we wanted the participant to have a better chance of learning the extensive number of items. Also, for the same reason we did not employ filler items. Second, the participant learned the material up to at least 60% (instead of 50%) correct in the training phase. Third, in the initial part of the TNT phase, the participant learned to recognize at least 90% (instead of 100%) of the cues as belonging to either think or no-think pairs, and did not engage in a practice session on the TNT task with the use of filler items. Fourth, in the critical part of the TNT phase, a response to a respond cue, as well as no response to a suppression cue, resulted in "correct" feedback. Such a feedback label was not used in the Anderson and Green study but was employed here to maintain motivation during the lengthy TNT phase. Fifth, in the test phase, cue words were shown for two seconds but the participant was allowed up to 30 s to respond (succeeded by a 5-s delay). In contrast, in Anderson and Green's study the exposure of cue words was response terminated or terminated after a latency of four seconds, whichever came first. Finally, for each participant items were randomly assigned to experimental conditions in all phases of the experiment.

3. Results

Data from the 48, respectively, 30 subjects are presented in Table 1. Also, Fig. 1 shows baseline-corrected recall (recall of items in the 8 and 16 repetition conditions after recall of items in the baseline has been subtracted out) as a function of repetitions with valence and TNT instructions as parameters. The pattern of results for the 48 participants is strikingly simple. Baseline recall (i.e., recall at 0-repetitions) was essentially the same irrespective of valence (negative = 77.9%, neutral = 80.4%, $p = .44$). As one might expect, recall in the think condition, in which the correct response was fedback whenever the participant made an error or could not respond, was well above baseline and roughly the same irrespective of valence. Recall of negative suppress items was similar to recall of negative baseline items suggesting that suppression of negative no-think items failed. In contrast, recall of neutral suppress items was well below recall of neutral baseline items suggesting that suppression of neutral no-think items succeeded. In addition, increasing the number of training cycles from 8 to 16 had negligible effects on recall.

All these effects were statistically reliable. First, an overall repeated measures analysis of variance revealed a significant main effect of TNT instructions, $F(1, 47) = 103.11$, $p < .001$, and of trial repetitions, $F(1, 47) = 3.58$, $p < .05$, a reliable interaction between TNT instructions and valence, $F(1, 47) = 14.59$, $p < .001$, a reliable interaction between TNT instructions and trial repetitions, $F(2, 94) = 36.09$, $p < .001$, but no other significant effects. Next, by a separate analysis of variance of the no-think condition, the effects of valence, repetitions, and the interaction between valence and repetitions were all significant (p's < .04). In the think condition, only trial repetitions had a significant effect ($p < .001$), while recall of both neutral and emotional words was about 94% and close to ceiling. In the no-think condition, there was a significant decrease in the recall of neutral items as a function of repetitions, $F(2, 94) = 8.28$, $p < .001$, but no such decrease for emotional items, $F(2, 94) < 1.0$. Finally, considering baseline-corrected data alone (i.e., only 8 and 16 repetitions), the difference between 8 and 16 repetitions was not significant irrespective of valence, and instructions, $F(1, 47) < 1.0$, for the main effect of repetitions as well as higher order interactions.

Turning to the results of the re-test we note that the pattern of data in the immediate recall test for the 30 participants that were re-tested was essentially the same as the pattern for the set of 48 participants (see Table 1). Baseline recall of negative and neutral items was not significantly different (negative = 74.7%, neutral = 82.7%, $p = .06$). Also, by averaging across 8 and 16 repetitions, which were not significantly different, and comparing with baseline data, a Bonferroni-corrected t-test revealed that recall of repeatedly rehearsed neutral and emotional targets was reliably above baseline, $p < 0.01$ and $p < 0.001$, respectively. Recall of items in the neutral suppress condition was reliably below baseline, $p < 0.02$, and recall of items in the emotional suppress condition was not significantly different from baseline, $p > 0.55$.

Baseline-corrected recall data for the 30 participants that served in the re-test are displayed in Fig. 2. Baseline recall in the

Table 1
Free recall of neutral and negative target words in the baseline (0 repetitions) as well as the suppress and respond conditions (8 and 16 repetitions).

	Condition (repetition)				
	Baseline (0)	Suppress (8)	Suppress (16)	Respond (8)	Respond (16)
Proportions correct free recall according to the condition (repetitions) at test and re-test for the neutral and negative words: means and (standard deviations)					
Neutral					
Test; $N = 48$	0.80 (0.20)	0.69 (0.26)	0.68 (0.24)	0.95 (0.09)	0.95 (0.08)
Test; $N = 30$	0.83 (0.20)	0.72 (0.24)	0.70 (0.21)	0.96 (0.07)	0.95 (0.08)
Re-test; $N = 30$	0.51 (0.27)	0.59 (0.25)	0.53 (0.22)	0.66 (0.18)	0.67 (0.26)
Negative					
Test; $N = 48$	0.78 (0.23)	0.77 (0.22)	0.77 (0.23)	0.94 (0.12)	0.94 (0.13)
Test; $N = 30$	0.75 (0.20)	0.79 (0.20)	0.75 (0.21)	0.93 (0.12)	0.93 (0.13)
Re-test; $N = 30$	0.47 (0.28)	0.53 (0.29)	0.55 (0.27)	0.60 (0.22)	0.67 (0.19)

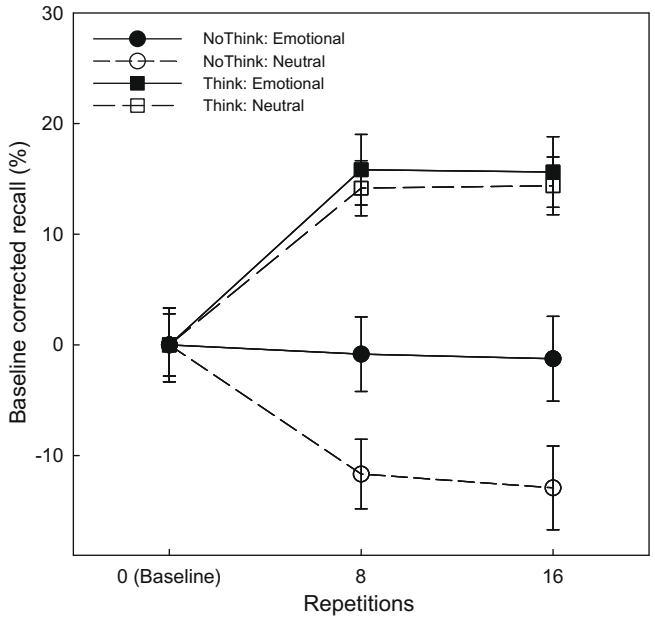

Fig. 1. Baseline-corrected recall of the neutral and emotional words in the suppress and respond conditions. Recall is shown as a function of number of repetitions during the think/no-think phase. Error bars indicate ±1 standard error of the mean.

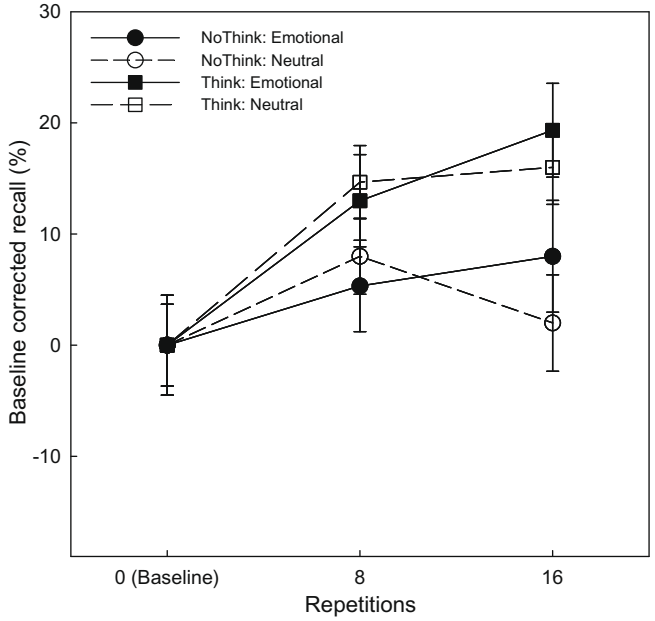

Fig. 2. Baseline-corrected recall of neutral and emotional words in the suppress and respond conditions after one week. Recall is shown according to the number of repetitions during the think/no-think phase. Error bars represent ±1 standard error of the mean.

re-test was roughly the same irrespective of valence (negative = 47.3%, neutral = 51.3%, $p = .52$). Again, the pattern of baseline-corrected data is pretty straightforward. Recall of repeatedly rehearsed items was well above baseline, whereas recall of neutral items previously attempted suppressed was similar to recall of emotional and baseline items. Thus, there was a significant effect of instructions (think/no-think), $F(1, 29) = 18.83$, $p < .001$, trial repetitions, $F(1, 29) = 12.08$, $p < .001$, and a significant interaction between instructions and repetitions, $F(2, 58) = 4.56$, $p < .02$. The interaction was due to a highly significant effect of repetitions in the think condition, $F(2, 58) = 20.76$, $p < .0001$, and no effect of repetitions in the no-think condition, $F(2, 58) = 1.85$, $p = .17$. The effect of repetitions (8 or 16) for baseline-corrected data paralleled the effects in the test and showed no reliable difference between 8 and 16 repetitions, and no significant interaction between repetitions, valence, and instructions.

Finally, direct comparisons of baseline-corrected recall between test and re-test data revealed a reliable difference on the to-be-suppressed neutral items by a Bonferroni-corrected t-test, ($p < .002$), but no other reliable differences between test and re-test.

4. Discussion

In agreement with prior studies, our findings suggest that consciously trying to forget previously encoded neutral words may succeed to some extent, but apparently not for a very long time, and in any case not for as long a period as a week. We suggest that the inferior recall of to-be-suppressed memories that can be measured immediately upon termination of no-think trials is a transient phenomenon, and note that the only study on volitional forgetting that to our knowledge has employed delayed tests (Wheeler, 1995) reported similar results, namely recovery of to-be-forgotten information during prolonged retention intervals.

Our findings on delayed retrieval may be interpreted within two different, but not necessarily incompatible, theoretical frameworks. Successfully suppressed material may have recovered due to a gradual release of inhibitory processes that were instantiated in the first place to keep particular targets out of mind. This would be an instance of relative spontaneous recovery because items in the suppress condition were characterized by a lower rate of forgetting during the delay compared to items in the baseline condition (see Brown, 1976 for a description of the distinction between relative and absolute spontaneous recovery). Below-baseline performance in volitional forgetting experiments is often explained with reference to inhibition (e.g., Anderson & Green, 2001; Bjork, 1989; Bjork et al., 2006; Levy & Anderson, 2002; but see Hertel & Calcaterra, 2005; MacLeod, 2007), i.e., the possibility that mental control causes the level of activation of a memory representation to be downregulated. If inhibition is intimately related to effortful and conscious control processes it makes sense that it is released when control is relinquished, when a person is no longer actively engaging in suppression. This interpretation is consistent with the view that inhibition is a momentary phenomenon (Bjork et al., 2006; MacLeod & Macrae, 2001; Wheeler, 1995; however, see Anderson, 2007).

However, it is suggestive that recall of items in the no-think condition was at least as good as (and actually slightly better than) recall of items in the baseline condition in the re-test (see Fig. 2). This is compatible with the notion that forgetting of baseline items was as fast as or faster than no-think items, and hence that instructions to not think of a subset of associates may have had the paradoxical effect of strengthening (compared to baseline) the memory of these very same associates. The only modern theory that directly addresses such contradictory findings is the ironic process theory of cognitive or mental control developed by Wegner and colleagues (e.g., Wegner, 2009; Wenzlaff & Wegner, 2000). This theory attempts to specify what we do when we try to suppress particular thoughts. When the task is to keep particular thoughts out of mind, the ironic process theory proposes that cognitive control mechanisms are configured such that one primarily conscious search process attempts to reach the desired mental state by filling the mind with unrelated thoughts, while another primarily unconscious process maintains the suppressed items in memory in order to control

that they are in fact kept out of mind. The operation of the latter process may ironically result in a preoccupation with the targets that are supposed to be avoided. On this account, then, baseline items should be forgotten at the same rate or faster than no-think items.

We found no evidence at all that emotional items were more susceptible to memory suppression than neutral items, let alone could be suppressed below baseline, as evidenced by the results from both the immediate and delayed re-tests (see Fig. 1 and 2). We suggest that this was the case because the emotional items were elaborated upon and encoded as more salient in memory compared to the neutral items. Existing research (e.g., Bywaters, Andrade, & Turpin, 2004; Kensinger & Corkin, 2003; Payne & Corrigan, 2007) indicates that emotional material is generally remembered better than neutral material, and also harder to suppress in thought suppression paradigms (for a review, see Wenzlaff & Wegner, 2000).

The notion that we are prone to forget or may in particular be unable to report negative material is fairly popular. It originates in psychoanalytic writings (see Anderson & Levy, 2002; Conway, 2001; Erdelyi, 2006), but has only recently been subjected to rigorous experimental tests. As previously noted, some studies do suggest that negative material may be subject to suppression (Depue et al., 2006, 2007; Joormann et al., 2005). We propose that the conflicting results stem from a minor but important difference in experimental design.[2] Depue et al. (2006, 2007) used a blocked design in which the impact of emotional words and pictures, which were all at a median level of arousal, may have leveled off after a few blocks with the result that both emotional and neutral stimuli were processed in a similar way. Also, Joormann et al. (2005) varied the valence of the suppress items between groups, and the participants may have segregated the suppress and respond words into sets according to the valence of the target words. In contrast, when emotional and neutral items are mixed, the appearance of emotional items is always unpredictable, and this may maintain emotional impact at a higher level. As a result, the suppression of emotional items should be harder than the suppression of neutral items. In fact, the findings of the one study which has employed a procedure similar to the present one and varied the valence of suppress (and respond) items within-subjects found that high arousal negative words were not suppressed (Marx et al., 2008).

Habituation may also explain why two studies which recently investigated list-method DF of neutral and emotional stimuli reported contrasting results. Using a blocked design, Wessel and Merckelbach (2006) found that DF was essentially the same for neutral and negative words, while Payne and Corrigan (2007) with a mixed design found that although DF occurred with respect to neutral pictures, memories of negative pictures resisted forgetting.

Although we consider these subtle methodological differences to be the most plausible reason why our results differ from some studies which have investigated suppression of negative material another possible explanation should be mentioned. One insidious problem with many studies which use emotional words is that such stimuli are often more closely associated with one another than neutral words (see Talmi & Moscovitch, 2004). Suppression of negative material may not have taken place in the present experiment because the words employed clustered together in semantic categories to a higher degree than in other experiments. However, this seems a remote possibility because we controlled for semantic relatedness by attempting to avoid that any words were related (see Appendix).

To sum up, we can control retrieval of neutral memories we do not want to think about if we try hard enough, but not completely and not for a very long time (certainly not for a week). So voluntary forgetting is temporary and to-be-suppressed items do recover. This presumably happens because they are only inhibited as long as no-think efforts are reinforced, or because the mechanisms set up for keeping the items out of mind at the same time at some level must keep track of these very same items to ensure that they do not come to mind. Can we suppress unpleasant memories? Recent research by Depue et al. (2006, 2007) and Joormann et al. (2005) suggests an affirmative answer, but our findings suggest that the amount of training that drives recall of neutral items below baseline has no measurable effect on recall of negative emotional items. An analysis of the conflicting results suggests that a difference in design is critical: Studies, in which the occurrence of negative emotional items is random, find no evidence that emotional material can be suppressed, while studies in which the incidence of emotional items is predictable do find such evidence. In conclusion, our results indicate that it is rather difficult to forget unwanted memories at will.

[2] It should be noted that the procedure employed by Depue et al. (2006, 2007) differed slightly from the one employed by Anderson and Green (2001) in that the former studies used equal numbers of think and no-think trials. This could have made it easier to suppress emotional material. Also, the studies by Depue et al. (2006, 2007) employed pictures as stimuli while we used words. This difference in the nature of stimuli material could be important. Finally, while habituation may explain the discordant results from different studies it is worth noting that we did not employ a control condition with blocked presentation of stimuli. Future studies should do so.

Acknowledgments

This research was supported by a grant from The Nordic Centre of Excellence on Cognitive Control. We thank Christian Gerlach and Kristin Munch Ryg for their helpful comments, Daniel Barratt and Johannes Lang for proofreading, Sander Van de Cruys and Mai Drost Nielsen for their help with the data collection, and Michael Anderson for kindly providing the questionnaire used in the study.

Appendix

Neutral and emotional words (and their English translations) employed in the experiment.

Neutral word pairs		Emotional word pairs	
Cue word	Target word	Cue word	Target word
Boks (Box)	Nøgleord (Keyword)	Akustik (Acoustics)	Gidsel (Hostage)
Brusebad (Shower)	Patent (Patent)	Albue (Elbow)	Foragt (Contempt)
Dusin (Dozen)	Vase (Vase)	Aspekt (Aspect)	Urin (Urine)
Forhæng (Curtain)	Fagområde (Subject area)	Blæk (Ink)	Nervøsitet (Nervousness)
Gøremål (Business)	Maling (Paint)	Bogstav (Letter)	Misundelse (Envy)
Håndtag (Handle)	Data (Data)	Dukke (Doll)	Dødsdom (Death sentence)
Import (Import)	Pensionat (Boarding house)	Etablering (Establishment)	Racist (Racist)
Juice (Juice)	Optakt (Prelude)	Foredrag (Lecture)	Bræk (Vomit)
Karosseri (Bodywork)	Indeks (Index)	Fotografi (Photography)	Leukæmi (Leukaemia)
Kompas (Compass)	Høring (Hearing)	Fundament (Foundation)	Rotte (Rat)
Lakrids (Liquorice)	Beføjelse (Authority)	Gødning (Fertilizer)	Virus (Virus)
Læder (Leather)	Fusion (Fusion)	Halvø (Peninsula)	Tortur (Torture)

Appendix (*continued*)

Neutral word pairs		Emotional word pairs	
Cue word	Target word	Cue word	Target word
Manøvre (Manoeuvre)	Pingvin (Penguin)	Indvielse (Opening)	Svulst (Tumour)
Mappe (Briefcase)	Layout (Layout)	Jungle (Jungle)	Incest (Incest)
Mikrofon (Microphone)	Opbevaring (Storage)	Kemi (Chemistry)	Røvhul (Arsehole)
Murer (Bricklayer)	Sektion (Section)	Kontrast (Contrast)	Dulle (Slut)
Norm (Norm)	Parabol (Satellite dish)	Krave (Collar)	Grusomhed (Cruelty)
Olympiade (Olympics)	Timian (Thyme)	Legat (Grant)	Sperm (Sperm)
Opførelse (Production)	Krat (Thicket)	Lineal (Ruler)	Fattigdom (Poverty)
Optælling (Counting)	Bagage (Luggage)	Lokomotiv (Locomotive)	Sindssyge (Insanity)
Piano (Piano)	Tidsalder (Era)	Læsning (Reading)	Parasit (Parasite)
Prognose (Forecast)	Grød (Porridge)	Ordsprog (Proverb)	Kusse (Cunt)
Prototype (Prototype)	Fregne (Freckle)	Overblik (Overview)	Idiot (Idiot)
Præcision (Precision)	Elefant (Elephant)	Pasta (Pasta)	Slaveri (Slavery)
Realisme (Realism)	Indgang (Entrance)	Spisetid (Mealtime)	Respirator (Respirator)
Ryglæn (Back)	Omlægning (Rearrangement)	Stav (Stick)	Dommedag (Doomsday)
Rytter (Rider)	Sendetid (Air time)	Svømmehal (Swimming bath)	Utroskab (Adultery)
Råolie (Crude oil)	Hårfarve (Colour of the hair)	Sømil (Nautical mile)	Satan (Satan)
Sagn (Legend)	Krølle (Curl)	Tabel (Table)	Ydmygelse (Humiliation)
Sengetøj (Bedding)	Råderum (Free scope)	Tankegang (Way of thinking)	Blodbad (Massacre)
Separat (Separate)	Indløb (Stream)	Tekstil (Textile)	Bombe (Bomb)
Topmøde (Summit)	Klap (Flap)	Terræn (Terrain)	Stress (Stress)
Trillebør (Wheelbarrow)	Etape (Stage)	Tiltag (Initiative)	Bøddel (Executioner)
Vane (Habit)	Balkon (Balcony)	Trilogi (Trilogy)	Asbest (Asbestos)
Ventil (Valve)	Kapacitet (Capacity)	Tømmer (Timber)	Tandpine (Toothache)

References

Andersen, M. S., Asmussen, H., & Asmussen, J. (2002). The project of korpus 2000 going public. In A. Braasch & C. Povlsen (Eds.), *Proceedings of the 10th EURALEX international congress* (pp. 291–299). Copenhagen: EURALEX 2002.

Anderson, M. C. (2007). Inhibition: Manifestations in long-term memory. In H. L. Roediger, Y. Dudai, & S. M. Fitzpatrick (Eds.), *Science of memory: Concepts* (pp. 295–299). New York: Oxford University Press.

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

Anderson, M. C., & Levy, B. (2002). Repression can (and should) be studied empirically. *Trends in Cognitive Sciences, 6*(12), 502–503.

Anderson, M. C., Ochsner, K. N., Kuhl, B., Cooper, J., Robertson, E., Gabrieli, S. W., et al. (2004). Neural systems underlying the suppression of unwanted memories. *Science, 303*, 232–235.

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

Bjork, R. A. (2001). Recency and recovery in human memory. In H. L. Roediger, J. S. Nairne, I. Neath, & A. M. Surprenant (Eds.), *The nature of remembering: Essays in honour of Robert G. Crowder* (pp. 211–232). USA: American Psychological Association.

Bjork, E. L., Bjork, R. A., & MacLeod, M. D. (2006). Types and consequences of forgetting: Intended and unintended. In L. G. Nilsson & N. Ohta (Eds.), *Memory and society: Psychological perspectives* (pp. 134–158). New York: Psychology Press.

Bradley, M. M., Lang, P. J., & Cuthbert, B. N. (1993). Emotion, novelty, and the startle reflex: Habituation in humans. *Behavioral Neuroscience, 107*(6), 970–980.

Brown, A. S. (1976). Spontaneous recovery in human learning. *Psychological Bulletin, 83*, 321–328.

Bulevich, J. B., Roediger, H. L., III., Balota, D. A., & Butler, A. C. (2006). Failures to find suppression of episodic memories in the think/no-think paradigm. *Memory and Cognition, 34*(8), 1569–1577.

Bywaters, M., Andrade, J., & Turpin, G. (2004). Determinants of the vividness of visual imagery: The effects of delayed recall, stimulus affect and individual differences. *Memory, 12*(4), 479–488.

Compton, R. J. (2003). The interface between emotion and attention: A review of evidence from psychology and neuroscience. *Behavioral and Cognitive Neuroscience Reviews, 2*, 115–129.

Conway, M. A. (2001). Repression revisited. *Nature, 410*, 319–320.

Depue, B. E., Banich, M. T., & Curran, T. (2006). Suppression of emotional and nonemotional content in memory – Effects of repetition on cognitive control. *Psychological Science, 17*(5), 441–447.

Depue, B. E., Curran, T., & Banich, M. T. (2007). Prefrontal regions orchestrate suppression of emotional memories via a two-phase process. *Science, 317*, 215–217.

Erdelyi, M. H. (2006). The unified theory of repression. *Behavioral and Brain Sciences, 29*, 499–551.

Feinstein, J. S., Goldin, P., Stein, M., Brown, G., & Paulus, M. (2002). Habituation of attentional networks during emotion processing. *NeuroReport, 13*(10), 1255–1258.

Gleaves, D. H., Smith, S. M., Butler, L. D., & Spiegel, D. (2004). False and recovered memories in the laboratory and clinic: A review of experimental and clinical evidence. *Clinical Psychology – Science and Practice, 11*(1), 3–28.

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

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

Joormann, J., Hertel, P. T., Brozovich, F., & Gotlib, I. H. (2005). Remembering the good, forgetting the bad: Intentional forgetting of emotional material in depression. *Journal of Abnormal Psychology, 114*(4), 640–648.

Kensinger, E. A., & Corkin, S. (2003). Memory enhancement for emotional words: Are emotional words more vividly remembered than neutral words? *Memory and Cognition, 31*(8), 1169–1180.

LaBar, K. S., & Cabeza, R. (2006). Cognitive neuroscience of emotional memory. *Nature Reviews Neuroscience, 7*, 54–64.

Levine, L. J., & Pizarro, D. A. (2004). Emotion and memory research: A grumpy overview. *Social Cognition, 22*(5), 530–554.

Levy, B. J., & Anderson, M. C. (2002). Inhibitory processes and the control of memory retrieval. *Trends in Cognitive Sciences, 6*(7), 299–305.

Levy, B. J., & Anderson, M. C. (2008). Individual differences in the suppression of unwanted memories: The executive deficit hypothesis. *Acta Psychologica, 127*, 623–635.

MacLeod, C. M. (1998). Directed forgetting. In J. M. Golding & C. M. MacLeod (Eds.), *Intentional forgetting: Interdisciplinary approaches* (pp. 1–57). Mahwah, NJ: Erlbaum.

MacLeod, C.M. (2007). Inhibition: Elusive or illusion? In H.L. Roediger, Y. Dudai, & S.M. Fitzpatrick (Eds.), *Science of memory: Concepts* (pp. 301–306). New York: Oxford University Press.

MacLeod, M. D., & Macrae, C. N. (2001). Gone but not forgotten: The transient nature of retrieval-induced forgetting. *Psychological Science, 12*, 148–152.

Marx, B.P., Marshall, P.J., & Castro, F. (2008).The moderating effects of stimulus valence and arousal on memory suppression. *Emotion, 8(2),* 199–207.

Payne, B. K., & Corrigan, E. (2007). Emotional constraints on intentional forgetting. *Journal of Experimental Social Psychology, 43*, 780–786.

Postman, L., Start, K., & Fraser, J. (1968). Temporal changes in interference. *Journal of Verbal Learning and Verbal Behavior, 7*, 672–694.

Reisberg, D., & Heuer, F. (2004). Memory for emotional events. In D. Reisberg & P. Hertel (Eds.), *Memory and emotion* (pp. 3–41). New York: Oxford University Press.

Talmi, D., & Moscovitch, M. (2004). Can semantic relatedness explain the enhancement of memory for emotional words? *Memory and Cognition, 32*(5), 742–751.

Wegner, D. M. (2009). How to think, say, or do precisely the worst thing for any occasion. *Science, 325*, 48–51.

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

Wessel, I., & Merckelbach, H. (2006). Forgetting 'murder' is not harder than forgetting 'circle': Listwise-directed forgetting of emotional words. *Cognition and Emotion, 20*(1), 129–137.

Wheeler, M. A. (1995). Improvement in recall over time without repeated testing: Spontaneous recovery revisited. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*, 173–184.

Wright, C. I., Fischer, H., Whalen, P. J., McInerney, S. C., Shin, L. M., & Rauch, S. L. (2001). Differential prefrontal cortex and amygdala habituation to repeatedly presented emotional stimuli. *NeuroReport, 12*, 379–383.