

Inducing and reducing false memories: A Swedish version of the Deese–Roediger–McDermott paradigm

MIKAEL JOHANSSON¹ and GEORG STENBERG^{1, 2}

¹Department of Psychology, Lund University, Sweden

²Department of Social Sciences, Växjö University, Sweden

Johansson, M. & Stenberg, G. (2002). Inducing and reducing false memories: A Swedish version of the Deese–Roediger–McDermott paradigm. *Scandinavian Journal of Psychology*, 43, 369–383.

Participants tend to falsely remember a nonpresented critical word after having studied a list of the word's primary associates. We present here a Swedish version of the Deese–Roediger–McDermott paradigm, which provides a tractable method of experimentally inducing and investigating such illusory memories. In Experiment 1 it was demonstrated that the constructed stimulus material induced highly reliable false-recall and false-recognition effects, and, moreover, that veridical and false memories were associated with a similar phenomenological experience of remembering. The results from Experiment 2 indicated that the susceptibility to false recognition can be substantially reduced when participants are explicitly required to monitor the sources of their memories. These findings are consistent with predictions derived from the source-monitoring framework.

Key words: False memories, memory illusion, recall, recognition, source monitoring.

Mikael Johansson, Department of Psychology, Lund University, PO Box 213, SE-221 00 Lund Sweden. E-mail mikael.johansson@psychology.lu.se

The study of memory illusions and memory distortions has attracted an increasing amount of interest and elucidated the constructive/reconstructive nature of human memory (see, e.g., Roediger, 1996; Roediger & McDermott, 2000; Schacter, 1995, 1996, for reviews). Distortions of remembering have been observed for a long time in a range of experimental paradigms, for example, the influence of schemas in recall and recognition (e.g., Bartlett, 1932; see Alba & Hasher, 1983, for a review), the effects of misleading suggestions (e.g., Loftus, Miller & Burns, 1978; see Loftus, 1993, for a review), and misattributions of processing fluency (e.g., Jacoby & Whitehouse, 1989; Whittlesea, 1993). Following the definition suggested by Roediger (1996), memory illusions are “cases in which a rememberer's report of a past event seriously deviates from the event's actual occurrence”.

The focus of the present study is the fact that people may under some circumstances believe that they have experienced an event that never actually happened. We present here a Swedish version of the widely used Deese–Roediger–McDermott (DRM; pronounced “dream”) paradigm for experimentally examining memory illusions in a standard list-learning procedure. Furthermore, we report a way of reducing the incidence of these compelling false memories by means of a source-monitoring procedure.

As originally reported by Deese (1959), participants tend to erroneously recall a nonpresented critical word (e.g., *sleep*) after having heard a list of the word's primary associates (e.g., *bed, rest, awake, tired, dream, wake, snooze, blanket, doze, slumber, snore, and nap*). Roediger and McDermott (1995) revived this paradigm and replicated and extended the original results. In their study, participants falsely recalled the critical nonpresented word in an immediate

recall test at levels comparable with the probability of accurate recall of words stemming from the middle of the list. Furthermore, in a subsequent recognition test, participants showed remarkably high levels of false recognition of the nonpresented lure words and, at the same time, very low levels of false alarms to unrelated distracter words. Interestingly, veridical recognition and false recognition were associated with equally high levels of confidence (Experiment 1) and “remember” responses (Experiment 2). The latter judgment was derived by the use of Tulving's (1985) *remember/know* procedure, in which participants are instructed to respond “remember” if they are able to recollect some specific aspect of the prior presentation episode of a given item and “know” if they are certain that an item was previously presented but fail to recollect contextual information about the event of the item's presentation. Thus, the results suggest that false recognition does not merely occur on the basis of familiarity but, rather, that veridical and false memories induced in this paradigm seem to share basic phenomenological qualities tapped by the “remember” judgment.

Since the introduction of the DRM paradigm, research has examined the influences of both encoding- and retrieval-related factors on veridical and illusory memories and the false-memory effect has proven to be a robust phenomenon (see Roediger, McDermott & Robinson, 1998, for a review). For example, even in situations in which participants are explicitly warned against making memory-judgment errors of this sort, false recognition remains high (Gallo, Roberts & Seamon, 1997; McDermott & Roediger, 1998). Equally remarkable is the finding that false recognition is still evident following extremely short encoding durations that leave veridical recognition at chance level (Seamon, Luo & Gallo, 1998).

A number of theoretical approaches have been proposed to account for the false-recall and false-recognition effects. One explanation refers to Underwood's (1965) *implicit associative response* (IAR). The idea is that the critical nonstudied item is consciously or unconsciously activated as a response to the encoding of list items during study. Then, in a subsequent recognition test, participants recognize the critical lure item but fail to remember the source of that activation (i.e., reality-monitoring failure; Johnson & Raye, 1981). Support for this explanation comes from results demonstrating a greater false-recognition effect when the study lists contain many, rather than few, associated words (Arndt & Hirshman, 1998; Robinson & Roediger, 1997), and the fact that higher levels of both false recall (McDermott, 1996) and false recognition (Mather, Henkel & Johnson, 1997; Tussing & Greene, 1997) are found following encoding of items grouped in lists than those found when study items are randomly mixed across lists.

While the above results are consistent with the IAR account, others are not as readily explained. For example, Payne, Elie, Blackwell and Neuschatz (1996) manipulated the retention interval and found differential effects on veridical and false recognition. While a significant decrease was evident in participants' correct recognition, the level of false recognition was unaffected when tested after a 24-hour delay. Similarly, McDermott (1996) observed a decrease in accurate recall but a slight increase in false recall over the same interval.

Further results that have been suggested as potential problems for the IAR account come from a slightly different experimental condition. In the studies reported by Koutstaal and Schacter (1997) and Seamon, Luo, Schlegel, Greene and Goldenberg (2000), participants studied pictures from different categories (e.g., cats, cars, shoes, etc.). In the subsequent recognition test, they falsely recognized nonstudied but category-related pictures. As argued by Koutstaal and Schacter, it seems unlikely that these new and highly distinctive pictures were generated as an automatic response during study. The authors therefore proposed that the false-recognition effect is due to the fact that the high degree of similarity among studied items leads to strong memory traces of what the related items have in common but weak memory traces of picture-specific details. Thus, new category-related items are falsely recognized because they correspond to the general theme of the studied items, an idea that is consistent with the following alternative account.

The *fuzzy-trace theory* (see Reyna & Brainerd, 1995, for a review) claims that memory judgments can be based on two different types of memory traces, namely gist and verbatim. While the former represent the meaning or theme of the event, the latter represent specific surface detail, such as the perceptual characteristics of the event. Veridical memory relies mainly on verbatim traces, whereas false recognition is based on gist traces (i.e., remembrance of the themes of the studied lists). Inherent in the theory is the notion that gist

and verbatim traces have different time courses and that memory for gist is more sustained in time than memory for verbatim information. Hence, the fuzzy-trace theory can, among other results, readily account for the differential effects of retention interval on veridical and false memory. However, as pointed out by Roediger *et al.* (1998; see also Schacter, Verfaellie & Pradere, 1996, for a similar remark) the theory has some difficulty in explaining why falsely recognized critical items attract "remember" responses. If false recognition is based on memory traces that represent gist information and that lack specific details about the study episode, one would expect these false memories to be given "know" responses.

Another potent account for the false-memory effects induced in the DRM paradigm is provided by the *source-monitoring framework* (SMF; Johnson, 1988; see Johnson, Hashtroudi & Lindsay, 1993; Mitchell & Johnson, 2000, for reviews). The SMF describes the processes involved in making attributions (and misattributions) about the source of remembered information and, hence, addresses the question of subjective experience of remembering directly. Briefly, the idea is that memories acquired from different sources differ in their distribution of various memory characteristics (e.g., perceptual, contextual, semantic, and affective information, and records of cognitive operations) and that these differences are evaluated by decision processes and used as cues for source-monitoring judgments. For example, memories for imagined events are typically associated with less perceptual and spatial information than memories for perceived events but contain instead relatively more information about reflective processes performed during encoding. Thus, source monitoring involves setting adequate criteria for the memory judgment at hand, for example, the level of familiarity (a high level implying that a stimulus was previously studied), or the amount of perceptual detail (a large amount suggesting that a stimulus was perceived rather than imagined). According to the SMF, remembered information can be assessed and evaluated rapidly and heuristically but also in a more strategic manner in which general knowledge is taken into account.

A number of findings support the notion that recognition and source monitoring require different amounts or types of information (e.g., Johansson, Stenberg, Lindgren & Rosén, 2002; Johnson, Kounios & Reeder, 1994). While almost any kind of stimulus information would be sufficient for correct recognition (e.g., perceptual fluency, semantic familiarity), more differentiated information is necessary for accurate source monitoring (e.g., color, form, spatial position, etc.). Hence, the SMF predicts that misattributions and false memories are more likely to occur when there is a great degree of overlap in memory characteristics associated with memories from different sources and, in addition, when a lenient criterion is being used (e.g., familiarity) compared with a more stringent criterion (e.g., perceptual information).

Turning to the DRM paradigm, it would be sufficient for participants to rely on the difference in semantic familiarity to discriminate well between studied words and unrelated distracter words. The apparent setback in using such a lenient criterion would be the resulting high level of false recognition of the critical items. In contrast, the SMF holds that false recognition should be reduced when encoding leads to differences in the memory characteristics associated with externally derived and internally generated memories and, furthermore, when these differences are monitored during remembering. In accordance with this notion, Smith and Hunt (1998) found a decrease in both false recall and false recognition when changing study presentation from auditory to visual modality. The authors argued that visual presentation promoted more distinctive item-specific encoding than did auditory presentation, which, in turn, made memories for visually presented items (relative to auditory items) more easily discriminated from internally generated memories.

Similarly, Israel and Schacter (1997) manipulated the distinctiveness of encoding by presenting each associate word on the lists auditorily and simultaneously displaying either a corresponding picture or a corresponding word. Participants in the pictorial-encoding condition exhibited a markedly reduced false-recognition rate compared with participants in the word-encoding condition. Israel and Schacter argued that the observed suppression of false memories was due to participants relying on a *distinctiveness heuristic* in the pictorial-encoding condition. The distinctiveness heuristic refers to participants' metamemorial awareness that veridical recognition should include recollection of distinctive details. That is, participants rejected critical lure items because they lacked access to pictorial information that was associated with memory for the studied items. Presumably, participants in the word-encoding condition did not demand this type of memory characteristic. Interpreted within the SMF, participants in the visual-encoding condition in the Smith and Hunt (1998) study and in the pictorial-encoding condition in the Israel and Schacter (1997) study adopted a more stringent criterion, querying their memories for perceptual information, which lowered the susceptibility to false memories.

In a similar vein, the second experiment reported in the present study investigated whether an explicit source-monitoring task would suppress false recognition in the DRM paradigm. The rationale for this was that accurate performance would require retrieval of source-specifying contextual information about the prior study episode and a reliance on semantic familiarity would be insufficient. Thus, participants would be encouraged to adopt a more stringent criterion relative to a recognition task and assess and evaluate memory characteristics revealing the nature of the study episode. Although critical lure items may be associated with high levels of semantic familiarity, source-specifying information should be absent or retrieved to a lesser degree and, as a consequence, lead to correct rejection of these nonstudied items.

The main objective of Experiment 1 was twofold: first, to construct Swedish stimulus material with the same specifications as the lists used in the DRM paradigm (i.e., lists containing the 15 strongest associates to a critical word and arranged in descending order of associative strength) and, second, to evaluate whether the developed lists would induce false memories analogous to those reported previously. For this latter purpose, Experiment 1 shared the basic design and procedure with that of Roediger and McDermott (1995, Experiment 2).

EXPERIMENT 1

Method

Participants. Sixty undergraduate students (39 women, 21 men) at the University of Lund volunteered to take part in the experiment. The mean age of the participants was 23.5 years (range 19–32) and all claimed to be native Swedish speakers. None had participated in the association test that was used to create the test lists (described below).

Materials. The lists were created in the same way as reported by Russell and Jenkins (1954) and used by Deese (1959) and Roediger and McDermott (1995). Two hundred and five participants (104 women, 101 men; mean age = 24.9, range 19–44) were instructed to give their first single-word association to each of 100 translated words from the Kent–Rosanoff Word Association Test. Participants' associations were then ranked with respect to the frequency of occurrence as a response. As a rule, when different inflected forms were given as associations, the most frequent one was selected and used. Thus, for each critical stimulus word, a list of semantically related words was obtained and arranged in descending order of associative strength. For our purposes, it was crucial to eliminate any overlap between lists (i.e., same word appearing as a response in more than one list). Because of this constraint, the number of lists was reduced to 48. On occasion, a few overlaps remained, but these were eliminated by removing the item from the list on which it was the less frequent. The final stimulus material is presented in the Appendix.

To conform to the procedure used by Roediger and McDermott (1995), the 48 lists were arbitrarily divided into two sets and tested in two different participant groups ($n = 30$ in each). For each group, the 24 lists were arbitrarily divided into three sets, which were used equally often in the three experimental conditions described below.

Three small booklets were created. The first contained eight blank pages that were used during free recall, the second multiplication and division problems, and the third the recognition test. Half the 96 items in the recognition test were studied and obtained by selecting items on serial positions 1, 8, and 10 from each studied list (48 targets). The other half were obtained by selecting items on serial positions 1, 8, and 10 from each nonstudied list (24 target controls) and by adding all 24 critical lure items (16 false targets and 8 false-target controls). These items were randomly ordered in the left column in the recognition-test booklet. The middle column displayed the response alternatives ("old", "new") and in the third column free space was provided for the remember/know judgment.

Design and procedure. The experiment involved a within-subjects design. Participants studied eight lists, and this was followed by an immediate free recall task, and then eight lists that were followed

by an arithmetic task. The remaining eight lists were not studied. Participants were tested individually or in small groups of no more than five.

Participants were told that they would hear a number of words and that their memory for these items would be tested. List items were presented in a male voice with the strongest associates occurring first, at a rate of approximately 1.5 s per item. After the 15th word in each list, the experimenter instructed participants to either recall the presented items or to work on the arithmetic problems. Participants were allowed 2 min for both these tasks before the next list began. Since the type of task was random across lists, participants did not know until after list presentation whether they would be required to recall the items or to work on the arithmetic problems. Participants wrote down all recalled items on a blank page and were told to close the booklet prior to the start of the subsequent list. Participants were, furthermore, instructed not to guess and told that they should be confident that all reported items were included in the list.

After the 16th list, participants were given written instructions for the old/new and the remember/know judgments, which were modeled on those employed by Rajaram (1993). Participants were thus told to respond "remember" if they could mentally re-experience the presentation of an item and respond "know" if contextual information of the item's presentation was lacking. In addition, they were explicitly told that the remember/know decision referred to the oral presentation and not to whether the item was produced during the free-recall task.

For each item in the recognition test, participants made an initial old/new judgment by circling either of the response alternatives "old" or "new" to mark their memory decision. Furthermore, for each recognized item, participants made their remember/know judgment by writing "R" or "K" before they continued with the next recognition test item.

After the recognition test, participants were debriefed. The complete experiment took approximately 1 hour.

Results and discussion

The reported analyses are based on data from all participants ($n = 60$) and therefore all of the 48 stimulus lists contributed to the results. If not otherwise noted, a p value of 0.05 was used for all statistical analyses in both the present and in the following experiment.

Recall

Studied targets were recalled at a mean rate of 0.53 and critical false targets at a mean rate of 0.43. As reported by Roediger and McDermott (1995), both primacy and recency effects were present in the data and the level of false recall was comparable to the level of accurate recall of list items in serial positions 4–11 (0.43), $t(59) < 1$, n.s. Crucially, intrusions other than the critical false target were made at a low rate (0.12). It should be noted that these intrusions were, with only a few exceptions, always words that were semantically congruent with the thematic content of the presented list. Thus, participants exhibited a reliable level of false recall, which was clearly higher than the proportion of intrusions of items other than the predicted false targets.

Recognition

Data from the recognition test are summarized in Table 1 and further illustrated in Fig. 1.

Performance measures were derived by subtracting the proportion of "old" responses to control items from "old" responses to studied targets and false targets, that is, $p(\text{targets}) = p(\text{"old" to targets}) - p(\text{"old" to target controls})$; $p(\text{false targets}) = p(\text{"old" to false targets}) - p(\text{"old" to false-target controls})$ (cf., high-threshold correction procedure, e.g., Schacter, Verfaellie, *et al.*, 1996). These measures were analyzed in a repeated measures analysis of variance (ANOVA), using item type (target vs. false target) and condition (recall vs. arithmetic) as factors. The main effect of condition was significant, $F(1, 59) = 8.58$, $MSE = 0.02$: there were more "old" responses to items from recall lists than to items on lists followed by the arithmetic task. Thus, having had the opportunity to recall items had a positive effect on later recognition performance for these items. Furthermore, the main effect of item type was found to be marginally significant, $F(1, 59) = 3.80$, $MSE = 0.02$, $p = 0.056$: participants tended to falsely recognize false targets to an even higher degree than they accurately recognized studied

Table 1. Proportion of "old" responses and performance measures for targets and false targets in Experiment 1

Item type and condition	Response alternative		Overall	Performance measure	
	Remember	Know		Old/new discrimination	Proportion of "remember"
<i>Targets</i>					
Study + recall	0.45 (0.02)	0.28 (0.02)	0.72 (0.02)	0.66 (0.02)	0.68 (0.04)
Study + arithmetic	0.32 (0.02)	0.34 (0.02)	0.65 (0.02)	0.59 (0.02)	0.53 (0.04)
Control	0.01 (0.00)	0.06 (0.01)	0.07 (0.01)		
<i>False targets</i>					
Study + recall	0.43 (0.03)	0.33 (0.02)	0.76 (0.03)	0.68 (0.02)	0.65 (0.05)
Study + arithmetic	0.34 (0.03)	0.39 (0.03)	0.72 (0.02)	0.64 (0.02)	0.55 (0.05)
Control	0.01 (0.00)	0.07 (0.02)	0.08 (0.02)		

Note: Standard errors are given in parentheses. Instances where "remember" and "know" responses do not add up to the overall performance reflect rounding error.

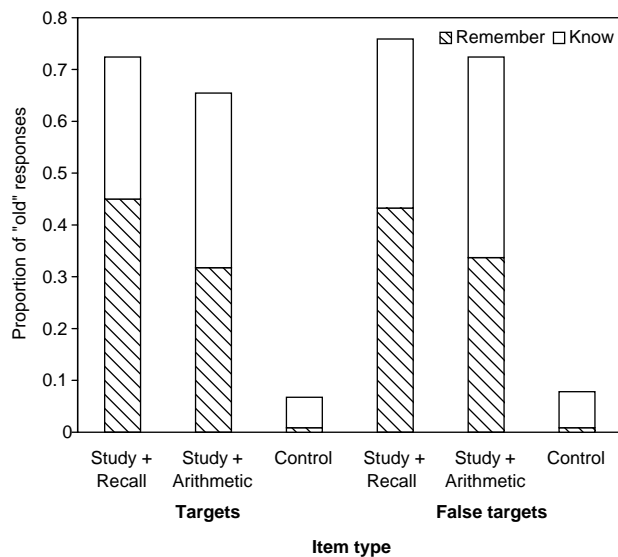


Fig. 1. Mean proportion of "remember" and "know" responses adding up to the overall proportion of "old" responses to targets and false targets in Experiment 1. Measures are shown separately for the three conditions used: study + recall, study + arithmetic, and control.

targets. The interaction between the two factors was not significant, $F(1, 59) = 1.03$, n.s.

As is suggested in Fig. 1, not only participants' recognition performance, but also their remember/know judgments were highly similar on targets and false targets. In addition, it can be seen that false alarms to target controls and false-target controls were at comparable levels and that they almost exclusively were accompanied by "know" responses. Since the proportion of remember and know responses add up to 1.0, we calculated a measure that reflected the proportion of the overall "old" responses associated with remember judgments, that is, $p(\text{remember}) = [p(\text{"remember"}) / p(\text{"remember"} + \text{"know"})]$, for each of the relevant experimental conditions after correction for false alarms to control items. These measures were subsequently subjected to a repeated measures ANOVA, employing the factors of item type (target vs. false target) and condition (recall vs. arithmetic). The only reliable effect found was the main effect of condition, $F(1, 59) = 21.78$, $MSE = 0.04$, which reflected the fact that participants made more "remember" responses to items from previously recalled lists. For all other effects, $F(1, 59) < 1$, n.s.

Included in the Appendix are the rates at which the critical lure items were falsely recalled and falsely recognized across participants. As can be seen, the lists produced false memories at different rates, ranging from 0% for *ilska* (anger) and *vissla* (whistle) to 90% for *flod* (river) and *hammare* (hammer) for false recall ($Mdn = 50\%$) and from 35% for *örn* (eagle) to 100% for *fönster* (window) and *ljus* (light) for false recognition ($Mdn = 75\%$). A similar variation

among lists has been found in previous research employing the lists included in the Roediger and McDermott (1995) study (Stadler, Roediger & McDermott, 1999). Stadler *et al.* observed, for example, low intrusion rates for the critical word *king* (10%), while *window*, *sleep*, *smell*, and *doctor* were falsely recalled at levels above 60%. In parallel, the false-recognition rates ranged from 27% for *king* to 84% for *window*, *sleep*, and *cold*. Although all lists were constructed in the same manner, they clearly vary in their effectiveness in producing false memories. Further research will examine list-specific factors that affect the false-memory rates and provide an account for the observed variations. The reason for giving the false-memory rates for each list in the Appendix is that these measures provide important information for selecting the material to use in subsequent research. In addition, such data highlight the importance of constructing the lists by means of an association test (as described in the Method section) and not by merely translating the list words from English to Swedish. For example, while Deese (1959) reported that the list constructed around the critical word *butterfly* failed to produce false recall, the same list was associated with an intrusion rate at 50% in the Swedish material.

In summary, the pattern of results found with the Swedish stimulus material presented here is in general agreement with that reported by Roediger and McDermott (1995), demonstrating robust false-recall and false-recognition effects in a list-learning paradigm. Moreover, not only was false recognition observed at a rate similar to veridical recognition, but the former tended to be even greater than the latter. In addition, the examination of participants' remember/know judgments shows that the false targets are phenomenologically similar to the studied targets, thus suggesting that the false-recognition effect is not merely mediated by a bias to respond "old" to all semantically familiar items and, further, that false and veridical memories induced in this procedure appear equally true to the participants.

EXPERIMENT 2

In the present experiment we tested the notion that the false-recognition effect may be reduced when participants are explicitly instructed to take source-specifying information into account, information that should be more difficult to find for the critical lure items. As was evident in Experiment 1, true and false memories are similar when the remember/know procedure is used to assess participants' phenomenology. However, the SMF holds that memories can differ across a wide range of characteristics. Therefore, it is possible that true and false memories are similar along the dimensions that are reflected in the general "remember" judgment, but differ instead in other aspects. Both Norman and Schacter (1997, Experiment 2) and Mather *et al.* (1997) investigated this issue by using modified versions of the Memory Characteristics Questionnaire (MCQ; Johnson,

Foley, Suengas & Raye, 1988) in addition to the remember/know procedure. The results indicated that true memories were associated with a higher amount of remembered perceptual (auditory) detail than were false memories for critical lure items but, in addition, that true and false memories were similar regarding memory for associations made when hearing the word. Moreover, Norman and Schacter (Experiment 1) asked participants to explain their "remember" responses by writing down what they recollected from the event of the items' presentation. For both false and veridical memories, these explanations contained mainly semantic/associative information and relatively few referred to perceptual/contextual information about the prior presentation.

Taken together, the above results suggest that there are subtle phenomenological differences between veridical and false memories and, furthermore, support the notion that participants' memory judgments may predominantly be based on information that is similar for targets and false targets. Consequently, it should be possible to lower the susceptibility to false memories in a task that instead requires participants to monitor average differences in qualitative characteristics. In the present experiment, we challenged the false-recognition effect by employing an immediate test after each list and by presenting items from two distinct sources, presumably yielding different memory characteristics and, hence, promoting source monitoring. Recently, Hicks and Marsh (1999) found that false recall was reduced in a condition in which participants had to decide whether recalled items were previously generated or heard, compared with a condition in which participants merely recalled previously heard words. We sought to examine whether this suppression effect extended to the domain of false recognition.

Method

Participants. Participants were 40 undergraduate students at the University of Lund (30 women, 10 men), who volunteered to take part in the study. The mean age of the participants was 22.7 years (range 19–32) and all claimed to be native Swedish speakers. Half of the participants were assigned to the old/new recognition task and the other half to the reality-monitoring task.

Materials. The 16 lists that produced the largest numbers of critical false alarms were selected from Experiment 1 and items on serial position 15 were dropped. The lists were divided into two sets for counter balancing purposes. The two sets had equal probability of inducing false recognition. Across subjects in each memory task, the list items were used equally often in the experimental conditions. Three types of small booklets were created. The first contained eight lined pages on which participants in the reality-monitoring task were to write down solutions to anagrams. The second contained basic multiplication and division problems for the 1-min retention interval. The third was the memory test and consisted of eight pages corresponding to the eight study lists. Each memory-test list comprised 18 items: one false target, one false-target control, eight targets (serial positions 1, 3, 4, 6, 7, 9, 10, and 12), and eight target

controls. The target controls were obtained by selecting one item from each nonstudied list (serial positions 1, 3, 4, 6, 7, 9, 10, and 12). The order of items was randomly determined for each memory-test page.

Design and procedure. The experiment involved a mixed factorial design. Participants were tested individually or in small groups of no more than five, and were told that their memory for words would be tested. The general procedure for all participants was as follows: first, they were presented with a list of 14 words; second, they were given 1 min to solve multiplication and division problems given in the math booklet; third, memory judgments were made on words in the memory-test booklet. Two minutes were allowed for the memory test, which was followed by the presentation of the next list. Detailed procedures for the two groups are described below.

In the old/new recognition task, participants heard list words presented in a male voice at approximately a 5-s rate. Following the arithmetic task, participants were instructed to turn to the designated page in the memory-test booklet. Test words were given in the left column and participants were to circle the response alternative "old" or "new" given in the right column to indicate their memory judgment.

In the reality-monitoring task, participants heard half of the words presented in a male voice and generated the other half as solutions to anagrams. The anagrams were presented as slides displayed for 3 s and participants were allowed an additional 2 s to write down the solution in the anagram booklet (see Hicks & Marsh, 1999; Marsh & Hicks, 1998, for a similar encoding task). Participants were instructed to write down the solution and then immediately cover the word with an opaque sheet of paper to ensure that past solutions were out of view. Thus, both generated and heard items were presented at a 5-s rate. Presentation of words alternated between generated and heard, beginning equally often with an anagram as with an auditorily presented word across lists. Following the arithmetic task, participants were instructed to turn to the designated page in the memory-test booklet. The response alternatives were "generated", "heard" and "new", and participants made their judgments by circling the alternative corresponding to their memory decision.

Following presentation of all eight lists, participants were debriefed. In all, the experiment took about 40 min in both the old/new recognition task and in the reality-monitoring task.

Results and discussion

The main results are displayed in Table 2. Since participants showed virtually no false alarms to either target controls or false-target controls, data for these item types are omitted from Table 2.

Recognition

A 2 (item type: target vs. false target) by 2 (task: old/new recognition vs. reality monitoring) mixed factor ANOVA was used to investigate participants' recognition performance. The results showed that there was a significant interaction between item type and task, $F(1, 38) = 26.75$, $MSE = 0.03$. To elucidate this interaction, follow-up analyses that compared performance across tasks were performed.

Two cross-task analyses were of primary interest, namely the comparison of memory performance to studied targets and the crucial comparison of the proportion of critical false alarms to false targets. As suggested in Table 2, participants in

Table 2. Proportion of "old" responses to targets and false targets in the two memory tasks used in Experiment 2

Item type	Old/new recognition	Reality monitoring		
		Both sources	Generated	Heard
<i>Targets</i>				
Overall	0.91 (0.02)	0.89 (0.01)		
For each source			0.91 (0.01)	0.88 (0.01)
<i>False targets</i>				
Overall	0.76 (0.05)	0.33 (0.05)		
Claimed source			0.06 (0.02)	0.28 (0.05)

Note: Standard errors are given in parentheses.

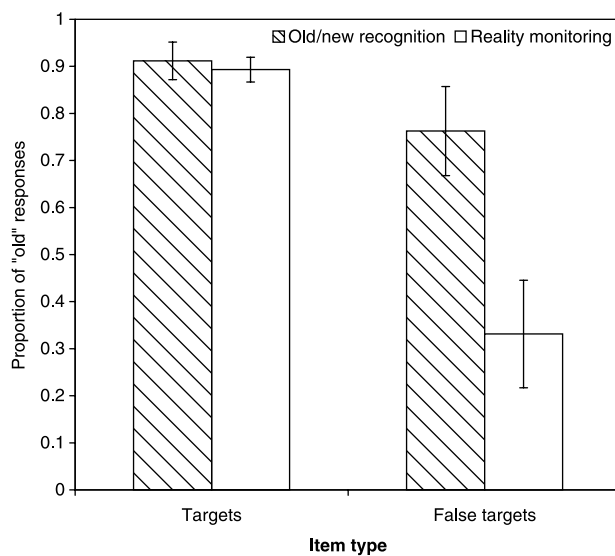


Fig. 2. Mean proportion of "old" responses to targets and false targets shown separately for the two memory tasks used in Experiment 2. The vertical lines depict confidence intervals (95%) for the means.

both the old/new recognition task and the reality-monitoring task accurately recognized previously studied targets at a high rate, and no difference was indeed found when contrasting the two memory tasks, $t(38) < 1.0$, n.s. Nonetheless, when turning to the critical false-alarm rate, the expected difference between tasks was found to be significant, $t(38) = 6.08$, $SED = 0.07$. As can be seen in Fig. 2, participants in the reality-monitoring task made fewer false alarms to semantically related lures than participants in the old/new recognition task, representing a reduction of 57%. Thus, the results are in accord with the idea that the explicit requirement to retrieve contextual information about the study episode reduces the susceptibility to false alarms of this kind.

Interestingly, the task manipulation lowered the number of critical false alarms, but it did not entirely eliminate the false-memory effect. The vertical lines in Fig. 2 depict the 95% confidence intervals and demonstrate that the effect was nevertheless reliable in the reality-monitoring task.

This pattern of results indicates the robustness of the false-memory effect induced in this paradigm.

Participants showed a higher level of recognition performance for generated items compared with heard items (0.91 vs. 0.88), $t(19) = 2.22$, $MSE = 0.01$. Albeit small, this generation effect parallels previous findings that show better memory performance with self-generated items than with items more passively encoded (e.g., Slamecka & Graf, 1978).

Source attribution

Data from the reality-monitoring task allowed an assessment of participants' source-attribution performance. Differences in source-monitoring performance related to previously generated and heard targets were assessed by comparing the proportion of erroneous source attributions associated with the two sources. The measures used were conditionalized on the overall level of recognition [i.e., $p(\text{wrong source}) / p(\text{correct source} + \text{wrong source})$], and indicated that participants misattributed more generated items than heard items (0.09 vs. 0.02), $t(19) = 3.83$, $MSE = 0.02$. That is, the different error rates suggest that participants were biased to respond "heard" when they failed to recollect the source of recognized targets.

A reliable false-memory effect was observed in the reality-monitoring task and, consequently, it was possible to examine participants' source attributions to the falsely remembered items. Such an examination would potentially give an indication of the phenomenological characteristics of participants' false memories. As can be seen in the last row of Table 2, participants exhibited a clear bias to judge their false memories as previously heard (0.28) rather than generated (0.06). This difference is significant, $t(19) = 4.41$, $MSE = 0.05$, and corresponds to the "it had to be you" effect described in source-monitoring research (e.g., Johnson, Raye, Foley & Foley, 1981). Different explanations of source-monitoring biases have been proposed. In the study by Hoffman (1997), participants attributed false alarms to the source associated with the lower recognition performance and the author argued that source-monitoring biases arise from average differences in memory strength for the two sources of items. This interpretation seems to be in accord

with the results of the present experiment, in which participants showed a slightly higher recognition performance to generated items than with heard items. However, the memory-strength account does not seem to be sufficient in explaining all source-monitoring biases. For example, Bink, Marsh and Hicks (1999) demonstrated that when recognition performance was equal for the two sources, participants misattributed new items to the source for which they showed the higher source accuracy. In explaining these source-monitoring biases, the authors argued that, depending on the test situation, different types of qualitative characteristics are considered more diagnostic. Presumably, participants in the present experiment considered information stemming from the generation procedure to be diagnostic (i.e., "if I had generated it, I would have remembered it") and consequently, if this type of contextual information is not recovered for a given item, but the item nonetheless feels familiar, participants believe that it was acquired from the other source (i.e., heard). Thus, both the memory-strength account and the referral to test-based decision criteria seem to be supported by the present results.

In summary, despite an immediate recognition test after each list, participants exhibited high levels of false recognition. This false-memory effect was, however, reduced in a task that required participants to find contextual information about the study event of each recognized item. Since the critical false targets were never presented, such information should be difficult to recover. This finding supports the notion that participants in the reality-monitoring task had to scrutinize the content of their memories more closely, which partly protected them from falsely recognizing critical lure items merely on the basis of semantic/associative information. Furthermore, since veridical recognition was equal in the two tasks, the reduction of false memories cannot be attributed to the use of a generally more conservative decision criterion in the reality-monitoring task. Thus, the task manipulation selectively affected the false-memory effect and left veridical memory performance unaffected.

GENERAL DISCUSSION

To summarize the main findings, we demonstrated in Experiment 1 that the Swedish version of the DRM paradigm presented here induced robust false-recall and false-recognition effects. False recall was observed at levels comparable to accurate recall of items from the middle of each list. In addition, participants exhibited false recognition of the critical items that tended to be even greater than correct recognition of the studied items. Furthermore, the immediate recall test increased the probability of both veridical recognition of list items and false recognition of critical nonstudied items. Like Roediger and McDermott (1995) and other authors, we employed the remember/know procedure to assess participants' phenomenology associated with true and false recognition. Using these measures, we

concluded that veridical and false recognition appear highly similar for the participants. However, previous results suggest that if participants are explicitly instructed to, they report a lower amount of perceptual detail for false memories than for true memories (Norman & Schacter, 1997; Mather *et al.*, 1997). As predicted by the SMF, the level of false recognition should be reduced if participants are required to capitalize on such differences during retrieval. In Experiment 2 we tested this notion by contrasting memory performance in a recognition task (which served as the baseline) with that in a reality-monitoring task. To promote source monitoring, we employed a two-source encoding task with the purpose of yielding distinct and different memory characteristics as a function of source. Presumably, memory for heard items would be associated with higher levels of auditory detail, whereas memory for generated items would be associated with a relatively higher amount of records of cognitive operations. As predicted, a reliable suppression of false recognition was indeed observed in the reality-monitoring task while the level of veridical recognition was equal in the two memory tasks. Thus, the present findings support and extend the results of Hicks and Marsh (1999) in showing that the susceptibility to false recognition in the DRM paradigm may be reduced by means of a source-monitoring procedure.

Taken together, the results observed in both of the present experiments indicate how compelling the false memories are in the DRM paradigm. For instance, despite the fact that memory performance was tested immediately after hearing a list of 15 words, participants showed a high level of false recognition of the critical nonstudied word. Likewise, although the explicit requirement to specify the origin of remembered information suppressed the false-recognition rate, the effect remained nonetheless significant. Thus, our attempts to challenge the false-recognition effect were not sufficiently powerful to eliminate the effect entirely.

As described in the Introduction, Israel and Schacter (1997) attributed their observed suppression effect to the use of a distinctiveness heuristic by which participants demand access to pictorial information before they endorse a test item as old. Apparently, their pictorial-encoding condition reduced false recognition in a test situation that does not require participants to take contextual information into account (i.e., old/new recognition). An answer to whether the suppression effect observed in Experiment 2 is due to a similar spontaneous and heuristic use of perceptual/contextual information awaits further research. Perhaps the hear/generate encoding condition is sufficient to make participants, in a parallel manner, adopt a more stringent criterion, or, alternatively, the explicit source-monitoring requirement is necessary.

Although elderly adults may sometimes be more prone to false remembering than young adults (e.g., Balota *et al.*, 1999; Cohen & Faulkner, 1989; Dywan & Jacoby, 1990; Koutstaal & Schacter, 1997; Norman & Schacter, 1997), recent results by Schacter, Israel and Racine (1999) indicate

that elderly adults show a suppression of false recognition after pictorial encoding, like young adults do. In contrast, Kensinger and Schacter (1999) showed that while young adults exhibit a reliable decrease in the false-recognition rate over repeated study/test trials (see McDermott, 1996, for a similar finding with false recall), elderly adults do not. The authors argued that elderly adults rely on memory for the general similarity of the studied items, but have difficulty in encoding or retrieving item-specific information. Following this reasoning, one explanation for the divergent results may be that the two tasks draw on different types of information. In the pictorial-encoding condition employed by Israel and Schacter, participants make use of distinctive qualities that are characteristic of all of the previously studied material (i.e., pictorial information), whereas the suppression effect found in the Kensinger and Schacter study is based on retrieval of item-specific information. Since the task employed in Experiment 2 above explicitly required participants to retrieve item-specific information to perform accurately, it would be interesting to examine whether elderly adults exhibit reduced levels of false recognition in this task as young adults apparently do.

Elderly adults' increased susceptibility to memory distortions has generally been attributed to source-memory deficits (see Spencer & Raz, 1995, for a review). Both neuropsychological data (e.g., Janowsky, Shimamura & Squire, 1989; Schacter, Harbluk & McClachlan, 1984; Shimamura & Squire, 1987) and results from neuroimaging studies (e.g., Rugg, Fletcher, Chua & Dolan, 1999; Senkfor & Van Petten, 1998; Wilding & Rugg, 1996) have implicated prefrontal regions in the retrieval of contextual information about the study episode. Only a few studies have examined false memories in the DRM paradigm with neuroimaging techniques. Interestingly, the results from two of these show increased anterior prefrontal activation during false compared with veridical recognition (see Schacter, Reiman, *et al.*, 1996, for positron emission tomography; see Schacter, Buckner, Koutstaal, Dale & Rosen, 1997, for functional magnetic resonance imaging). Schacter, Reiman, *et al.* (1996) suggested that this increased frontal-lobe activity reflects postretrieval processes that attempt to scrutinize recollective experience associated with the critical nonstudied items. Thus, as suggested by Norman and Schacter (1997), prefrontal dysfunction due to aging or damage would pose difficulties in resisting illusory memories of the critical items. However, further research is needed to directly link the mechanisms of the false-memory suppression effect to the prefrontal regions.

While both the IAR and the SMF accounts explain false remembering in the DRM paradigm by referring to source confusions, only the SMF is sufficiently specified as to allow predictions concerning factors at both encoding and retrieval that influence the occurrence and nature of these false memories. As noted in the Introduction, it is not apparent how the fuzzy-trace theory, which can readily

accommodate a large number of previously reported findings, explains the fact that participants in the DRM paradigm claim to vividly "remember" the critical nonstudied items. One possibility is that gist-based false memories attract "remember" judgments due to retrieval of verbatim information for other study-list items. However, it is not clear how such misattribution processes operate, nor is it clear how the theory explains the finding in Experiment 2 that participants showed a clear bias to claim that they had heard rather than generated the falsely remembered critical items.

Although it is beyond the scope of the present study to give an extensive review and evaluation of the various theoretical accounts proposed for the false memories induced in the DRM paradigm, we conclude that the results of the present study are consistent with predictions derived from the SMF. Furthermore, the SMF provides a potential account for a wide range of other memory distortions, for example, eyewitness suggestibility (e.g., Loftus, 1993), false-fame judgments (Jacoby, Woloshyn & Kelley, 1989), cryptomnesia (e.g., Marsh & Bower, 1993; Marsh, Landau & Hicks, 1997), and illusions of truth (Begg, Anas & Farinacci, 1992). Analogously to the suppression effect observed in Experiment 2 above, Lindsay and Johnson (1989) reported that the eyewitness-suggestibility effect was reduced in a source-monitoring task compared with a recognition task in which participants reliably misattributed memories of misleading suggestions as memories from the witnessed event. Thus, instructing participants to monitor the sources of their memories may reduce the susceptibility to other false-memory phenomena as well.

In summary, the Swedish version of the DRM paradigm presented here induced robust false-recall and false-recognition effects comparable to those previously reported (Roediger & McDermott, 1995). Furthermore, an explicit source-monitoring task reliably reduced the proportion of falsely recognized critical lure items. Because of the remarkably high levels of false memories observed in the DRM paradigm and the similarity in phenomenology associated with veridical and false remembering, it constitutes an appealing way of investigating the processes responsible for both accurate and distorted memories as well as factors influencing the susceptibility to memory illusions.

This research was supported by the Swedish Council for Research in the Humanities and Social Sciences (HSFR #F0763).

REFERENCES

- Alba, J. W. & Hasher, L. (1983). Is memory schematic? *Psychological Bulletin*, *93*, 203–231.
- Arndt, J. & Hirshman, E. (1998). True and false recognition in MINERVA2: Explanations from a global matching perspective. *Journal of Memory and Language*, *39*, 371–391.
- Balota, D. A., Cortese, M. J., Duchek, J. M., Adams, D., Roediger, H. L., McDermott, K. B. & Yerys, B. E. (1999). Veridical and false memories in healthy older adults and in dementia of the Alzheimer's type. *Cognitive Neuropsychology*, *16*, 361–384.

- Bartlett, F. C. (1932). *Remembering: A study in experimental and social psychology*. Cambridge: Cambridge University Press.
- Begg, I. M., Anas, A. & Farinacci, S. (1992). Dissociation of processes in belief: Source recollection, statement familiarity, and the illusion of truth. *Journal of Experimental Psychology: General*, *121*, 446–458.
- Bink, M. L., Marsh, R. L. & Hicks, J. L. (1999). An alternative conceptualization to memory “strength” in reality monitoring. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *25*, 804–809.
- Cohen, G. & Faulkner, D. (1989). Age differences in source forgetting: Effects on reality monitoring and on eyewitness testimony. *Psychology and Aging*, *4*, 10–17.
- Deese, J. (1959). On the prediction of occurrence of particular verbal intrusions in immediate recall. *Journal of Experimental Psychology*, *58*, 17–22.
- Dywan, J. & Jacoby, L. (1990). Effects of aging on source monitoring: Differences in susceptibility to false fame. *Psychology and Aging*, *5*, 379–387.
- Gallo, D. A., Roberts, M. J. & Seamon, J. G. (1997). Remembering words not presented in lists: Can we avoid creating false memories? *Psychonomic Bulletin and Review*, *4*, 271–276.
- Hicks, J. L. & Marsh, R. L. (1999). Attempts to reduce the incidence of false recall with source monitoring. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *25*, 1195–1209.
- Hoffman, H. G. (1997). Role of memory strength in reality monitoring decisions: Evidence from source attribution biases. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *23*, 371–383.
- Israel, L. & Schacter, D. L. (1997). Pictorial encoding reduces false recognition of semantic associates. *Psychonomic Bulletin and Review*, *4*, 577–581.
- Jacoby, L. L. & Whitehouse, K. (1989). An illusion of memory: False recognition influenced by unconscious perception. *Journal of Experimental Psychology: General*, *118*, 126–135.
- Jacoby, L. L., Woloshyn, V. & Kelley, C. (1989). Becoming famous without being recognized: Unconscious influences of memory produced by dividing attention. *Journal of Experimental Psychology: General*, *118*, 115–125.
- Janowsky, J. S., Shimamura, A. P. & Squire, L. R. (1989). Source memory impairment in patients with frontal lobe lesions. *Neuropsychologia*, *27*, 1043–1056.
- Johansson, M., Stenberg, G., Lindgren, M. & Rosén, I. (2002). Memory for perceived and imagined pictures: An event-related potential study. *Neuropsychologia*, *90*, 986–1002.
- Johnson, M. K. (1988). Reality monitoring: An experimental phenomenological approach. *Journal of Experimental Psychology: General*, *117*, 390–394.
- Johnson, M. K., Foley, M. A., Suengas, A. G. & Raye, C. L. (1988). Phenomenal characteristics of memories for perceived and imagined autobiographical events. *Journal of Experimental Psychology: General*, *117*, 371–376.
- Johnson, M. K., Hashtroudi, S. & Lindsay, D. S. (1993). Source monitoring. *Psychological Bulletin*, *114*, 3–28.
- Johnson, M. K., Kounios, J. & Reeder, J. A. (1994). Time-course studies of reality monitoring and recognition. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *20*, 1409–1419.
- Johnson, M. K. & Raye, C. L. (1981). Reality monitoring. *Psychological Review*, *88*, 67–85.
- Johnson, M. K., Raye, C. L., Foley, H. J. & Foley, M. A. (1981). Cognitive operations and decision bias in reality monitoring. *American Journal of Psychology*, *94*, 37–64.
- Kensinger, E. A. & Schacter, D. L. (1999). When true memories suppress false memories: Effects of ageing. *Cognitive Neuropsychology*, *16*, 399–415.
- Koutstaal, W. & Schacter, D. L. (1997). Gist-based false recognition of pictures in older and younger adults. *Journal of Memory and Language*, *37*, 555–583.
- Lindsay, D. S. & Johnson, M. K. (1989). The eyewitness suggestibility effect and memory for source. *Memory and Cognition*, *17*, 349–358.
- Loftus, E. F. (1993). Made in memory: Distortions in recollection after misleading information. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in theory and research* (vol. 27, pp. 187–215). New York: Academic Press.
- Loftus, E. F., Miller, D. G. & Burns, H. J. (1978). Semantic integration of verbal information into a visual memory. *Journal of Experimental Psychology: Human Learning and Memory*, *4*, 19–31.
- Marsh, R. L. & Bower, G. H. (1993). Eliciting cryptomnesia: Unconscious plagiarism in a puzzle task. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *19*, 673–688.
- Marsh, R. L. & Hicks, J. L. (1998). Test formats change source-monitoring decision processes. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *22*, 1296–1303.
- Marsh, R. L., Landau, J. D. & Hicks, J. L. (1997). Contributions of inadequate source monitoring to unconscious plagiarism during idea generation. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *23*, 886–897.
- Mather, M., Henkel, L. A. & Johnson, M. K. (1997). Evaluating characteristics of false memories: Remember/know judgments and memory characteristics questionnaire compared. *Memory and Cognition*, *25*, 826–837.
- McDermott, K. B. (1996). The persistence of false memories in list recall. *Journal of Memory and Language*, *35*, 212–230.
- McDermott, K. B. & Roediger, H. L. (1998). Attempting to avoid illusory memories: Robust false recognition of associates persists under conditions of explicit warnings and immediate testing. *Journal of Memory and Language*, *39*, 508–520.
- Mitchell, K. J. & Johnson, M. K. (2000). Source monitoring: Attributing mental experiences. In E. Tulving & F. I. M. Craik (Eds.), *The Oxford handbook of memory* (pp. 179–195). New York: Oxford University Press.
- Norman, K. A. & Schacter, D. L. (1997). False recognition in younger and older adults: Exploring the characteristics of illusory memories. *Memory and Cognition*, *25*, 838–848.
- Payne, D. G., Elie, C. J., Blackwell, J. M. & Neuschatz, J. S. (1996). Memory illusions: Recalling, recognizing, and recollecting events that never occurred. *Journal of Memory and Language*, *35*, 261–285.
- Rajaram, S. (1993). Remembering and knowing: Two means of access to the personal past. *Memory and Cognition*, *21*, 89–102.
- Reyna, V. F. & Brainerd, C. J. (1995). Fuzzy-trace theory: An interim synthesis. *Learning and Individual Differences*, *7*, 1–75.
- Robinson, K. J. & Roediger, H. L. (1997). Associative processes in false recall and false recognition. *Psychological Science*, *8*, 231–237.
- Roediger, H. L. (1996). Memory illusions. *Journal of Memory and Language*, *35*, 76–100.
- Roediger, H. L. & McDermott, K. B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *21*, 803–814.
- Roediger, H. L. & McDermott, K. B. (2000). Distortions of memory. In E. Tulving & F. I. M. Craik (Eds.), *The Oxford handbook of memory* (pp. 149–162). New York: Oxford University Press.
- Roediger, H. L., McDermott, K. B. & Robinson, K. J. (1998). The role of associative processes in creating false memories. In M. A. Conway, S. E. Gathercole & C. Cornoldi (Eds.), *Theories of memory II* (pp. 187–245). Hove, Sussex: Psychological Press.

- Rugg, M. D., Fletcher, P. C., Chua, P. M. & Dolan, R. J. (1999). The role of the prefrontal cortex in recognition memory and memory for source: An fMRI study. *Neuroimage*, *10*, 520–529.
- Russell, W. A. & Jenkins, J. J. (1954). *The complete Minnesota norms for responses to 100 words from the Kent–Rosanoff Word Association Test* (Tech. Rep. No. 11, Contract N8 ONR 66216), University of Minnesota, Office of Naval Research.
- Schacter, D. L. (1995). Memory distortions: History and current status. In D. L. Schacter, J. T. Coyle, G. D. Fischbach, M.-M. Mesulam & L. E. Sullivan (Eds.), *Memory distortion: How minds, brains and societies reconstruct the past* (pp. 1–43). Cambridge, MA: Harvard University Press.
- Schacter, D. L. (1996). *Searching for memory: The brain the mind and the past*. New York: Basic Books.
- Schacter, D. L., Buckner, R. L., Koutstaal, W., Dale, A. M. & Rosen, B. R. (1997). Late onset of anterior prefrontal activity during true and false recognition: An event-related fMRI study. *Neuroimage*, *6*, 259–269.
- Schacter, D. L., Harbluk, J. L. & McClachlan, D. R. (1984). Retrieval without recollection: An experimental analysis of source amnesia. *Journal of Verbal Learning and Verbal Behavior*, *23*, 593–611.
- Schacter, D. L., Israel, L. & Racine, C. (1999). Suppressing false recognition in younger and older adults: The distinctiveness heuristic. *Journal of Memory and Language*, *40*, 1–24.
- Schacter, D. L., Reiman, E., Curran, T., Yun, L. S., Bandy, D., McDermott, K. B. & Roediger, H. L. (1996). Neuroanatomical correlates of veridical and illusory recognition memory: Evidence from positron emission tomography. *Neuron*, *17*, 267–274.
- Schacter, D. L., Verfaellie, M. & Pradere, D. (1996). The neuropsychology of memory illusions: False recall and recognition in amnesic patients. *Journal of Memory and Language*, *35*, 319–334.
- Seamon, J. G., Luo, C. R. & Gallo, D. A. (1998). Creating false memories of words with or without recognition of list items: Evidence for nonconscious processes. *Psychological Science*, *9*, 20–26.
- Seamon, J. G., Luo, C. R., Schlegel, S. E., Greene, S. E. & Goldenberg, A. B. (2000). False memory for categorized pictures and words: The category associates procedure for studying memory errors in children and adults. *Journal of Memory and Language*, *42*, 120–146.
- Senkfor, A. J. & Van Petten, C. (1998). Who said what? An event-related potential investigation of source and item memory. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *24*, 1005–1025.
- Shimamura, A. P. & Squire, L. R. (1987). A neuropsychological study of fact memory and source amnesia. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *13*, 464–473.
- Slamecka, N. J. & Graf, P. (1978). The generation effect: Delineation of a phenomenon. *Journal of Experimental Psychology: Human Learning and Memory*, *4*, 592–604.
- Smith, R. E. & Hunt, R. R. (1998). Presentation modality affects false memory. *Psychonomic Bulletin and Review*, *5*, 710–715.
- Spencer, W. D. & Raz, N. R. (1995). Differential effects of aging on memory for content and context: A meta-analysis. *Psychology and Aging*, *10*, 527–539.
- Stadler, M. A., Roediger, H. L. & McDermott, K. B. (1999). Norms for word lists that create false memories. *Memory and Cognition*, *27*, 494–500.
- Tulving, E. (1985). Memory and consciousness. *Canadian Psychologist*, *21*, 1–12.
- Tussing, A. A. & Greene, R. L. (1997). False recognition of associates: How robust is the effect? *Psychonomic Bulletin and Review*, *4*, 572–576.
- Underwood, B. J. (1965). False recognition produced by implicit verbal responses. *Journal of Experimental Psychology*, *70*, 122–129.
- Whittlesea, B. W. A. (1993). Illusions of familiarity. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *19*, 1235–1253.
- Wilding, E. L. & Rugg, M. D. (1996). An event-related potential study of recognition memory with and without retrieval of source. *Brain*, *119*, 889–905.

Received 27 November 2000, accepted 6 June 2001

APPENDIX. THE 48 15-WORD LISTS AND THEIR CRITICAL ITEMS

List words are arranged in descending order of associative strength to the critical lure word. The rates at which each critical lure was falsely remembered are given in parentheses (recall/recognition).

<i>Arbete</i> (0.60/0.90)	<i>Bad</i> (0.10/0.75)	<i>Barn</i> (0.40/0.80)	<i>Befäl</i> (0.10/0.80)
Jobb	Kar	Små	Militär
Pengar	Ren	Leka	Soldat
Slita	Hav	Vuxna	Officer
Fritid	Skum	Dagis	Order
Ledig	Vått	Baby	Chef
Lön	Simma	Framtid	Uniform
Inkomst	Bubblor	Mamma	Överste
Kontor	Dusch	Familj	Armé
Semester	Smutsig	Syskon	Kapten
Fabrik	Bastu	Joller	Lumpen
Dator	Brygga	Napp	Auktoritär
Sysselsättning	Pool	Vagn	Menig
Rast	Klor	Gravid	Furir
Utbildning	Tvål	Föräldrar	Kommando
Anställning	Bassäng	Födelse	Honnör
<i>Berg</i> (0.50/0.60)	<i>Blomma</i> (0.60/0.75)	<i>Bröd</i> (0.50/0.60)	<i>Doktor</i> (0.10/0.60)
Dal	Röd	Mat	Sjukhus
Snö	Ros	Smör	Läkare
Topp	Blad	Äta	Vitklädd
Alper	Växt	Skiva	Stetoskop
Sten	Stjälk	Frukost	Rock
Klättra	Bi	Limpa	Patient
Höjd	Äng	Macka	Sköterska
Klippa	Träd	Kniv	System
Skidor	Lukta	Rosta	Medicin
Fjäll	Tulpan	Baka	Recept
Norge	Vas	Smulor	Kirurg
Granit	Bukett	Torrt	Hälsa
Stup	Flora	Fullkorn	Klinik
Gruva	Kruka	Vete	Väntrum
Kebnekaise	Eternell	Ost	Återbesök
<i>Fjäril</i> (0.50/0.80)	<i>Flicka</i> (0.50/0.90)	<i>Flod</i> (0.90/0.95)	<i>Fot</i> (0.50/0.70)
Sommar	Pojke	Vatten	Tå
Vingar	Flätor	Ebb	Hand
Puppa	Tös	Å	Skor
Larv	Ung	Nilen	Kroppsdel
Fladdrar	Tjej	Bäck	Nagel
Mask	Kjol	Fors	Svett
Skör	Klänning	Fisk	Vårta
Vår	Oskuld	Ström	Strumpa
Håv	Hopprep	Älv	Ankel
Kokong	Dockor	Ganges	Häl
Mygga	Dotter	Översvämning	Boll
Silke	Fräknar	Sjö	Sula
Papillon	Fästmö	Amazonas	Arm
Slända	Jänta	Fjord	Sparka
Fånga	Piga	Flotte	Sandal

<i>Frukt</i> (0.20/0.60)	<i>Fönster</i> (0.50/1.00)	<i>Gata</i> (0.40/0.80)	<i>Grov</i> (0.50/0.90)
Äpple	Glas	Väg	Fin
Apelsin	Utsikt	Stad	Stor
Banan	Hus	Kullersten	Skrovlig
Grönsak	Gardin	Asfalt	Muskulös
Päron	Ruta	Bilar	Sandpapper
Bär	Dörr	Trottoar	Rå
Nyttigt	Karm	Hemma	Slät
Korg	Putsa	Rak	Kraftig
Citron	Genomskinligt	Lykta	Sträv
Juice	Luft	Aveny	Tunn
Kiwi	Öppet	Nummer	Rejäl
Exotisk	Spröjs	Trafik	Späd
Förbjuden	Hasp	Gränd	Robust
Saftig	Persienn	Körbana	Vulgär
Mogen	Markis	Esplanad	Brutal
<i>Hammare</i> (0.90/0.85)	<i>Hög</i> (0.40/0.75)	<i>Ilska</i> (0.00/0.65)	<i>Jorden</i> (0.80/0.70)
Spik	Låg	Arg	Rund
Verktyg	Droger	Glädje	Månen
Slå	Knark	Vrede	Planet
Tor	Torn	Frustration	Klot
Snickardon	Skyskrapa	Hat	Tellus
Skruvmejsel	Kulle	Aggression	Universum
Tumme	Eiffeltornet	Förbannad	Världen
Såg	Gräs	Blixt	Glob
Slöjd	Marijuana	Raseri	Himlen
Stål	Påtänd	Smärta	Brun
Städ	Stege	Skratta	Rymden
Öra	Svindlande	Temperament	Miljö
Banka	Kran	Adrenalin	Snurrar
Yxa	Lyfta	Gräl	Resurs
Skära	Upp	Vansinne	Plöja
<i>Kall</i> (0.30/0.70)	<i>Kung</i> (0.30/0.75)	<i>Kvadrat</i> (0.50/0.95)	<i>Lejon</i> (0.10/0.65)
Varm	Drottning	Fyrkant	Gul
Is	Krona	Matte	Tiger
Vinter	Monarki	Cirkel	Farlig
Frysa	Slott	Geometri	Afrika
Kylig	Sverige	Kub	Savann
Glass	Tron	Figur	Stjärntecken
Sval	Makt	Rot	Djungel
Stel	Härskare	Triangel	Gap
Frost	Konung	Rektangel	Horoskop
Termometer	Fosterland	Form	Vilda
Vind	Kejsare	Låda	Ryta
Blöt	Regent	Meter	Simba
Okänslig	Rike	Torg	Zoo
Huttra	Prins	Romb	Antilop
Väder	Baron	Fyrhörning	Cirkus

<i>Ljus</i> (0.80/1.00)	<i>Lång</i> (0.80/0.75)	<i>Långsam</i> (0.50/0.75)	<i>Man</i> (0.10/0.60)
Mörk	Kort	Snabb	Kvinna
Stearin	Smal	Snigel	Skägg
Lampa	Flaggstång	Slö	Karl
Sol	Orm	Trög	Kille
Blond	Stång	Seg	Stark
Dag	Ståtlig	Sakta	Stilig
Värme	Sträcka	Tråkig	Kärlek
Advent	Avstånd	Sköldpadda	Kön
Låga	Reslig	Fort	Macho
Sken	Gänglig	Lat	Kostym
Tändsticka	Basketspelare	Sölig	Person
Levande	Linjal	Promenad	Knöl
Veke	Tunnel	Utdragen	Make
Brinna	Giraff	Väntan	Farbror
Paraffin	Korridor	Försiktig	Gubbe
<i>Matta</i> (0.60/0.85)	<i>Mjuk</i> (0.40/0.85)	<i>Musik</i> (0.50/0.65)	<i>Nål</i> (0.60/0.90)
Golv	Hård	Toner	Tråd
Frans	Kudde	Noter	Sy
Rya	Len	Gitarr	Vass
Persisk	Nalle	Sång	Stick
Dammsuga	Filt	Instrument	Öga
Trasa	Kramdjur	Mozart	Spruta
Randig	Bomull	Dans	Blod
Hall	Päls	Radio	Hål
Äkta	Gosig	Lyssna	Höstack
Tvätta	Ull	Ljuv	Akupunktur
Orientalisk	Moln	Klassisk	Spetsig
Ombonat	Fluffig	Piano	Häftstift
Plankor	Kind	Symfoni	Kanyl
Plast	Plysch	Konsert	Virka
Underlag	Frotté	Orkester	Fingerborg
<i>Präst</i> (0.30/0.75)	<i>Rädd</i> (0.60/0.90)	<i>Rättvisa</i> (0.40/0.70)	<i>Spindel</i> (0.20/0.50)
Kyrka	Skrämd	Domstol	Nät
Krage	Ensam	Lag	Ben
Bibel	Modig	Fred	Äcklig
Gud	Skraj	Demokrati	Hårig
Kappa	Feg	Juridik	Kryp
Kors	Trygg	Frihet	Insekt
Religion	Ledsen	Lika	Fobi
Pastor	Ängslig	Socialism	Väv
Vigsel	Osäker	Advokat	Åtta
Bröllop	Darra	Samhälle	Luden
Dop	Orolig	Jämställt	Läskig
Konfirmation	Obehagligt	Ordning	Fluga
Predika	Fly	Etik	Giftig
Munk	Monster	Diskriminering	Tarantella
Påven	Tapper	Fördelning	Änka

<i>Stol</i> (0.60/0.75)	<i>Svart</i> (0.70/0.55)	<i>Sömn</i> (0.50/0.70)	<i>Söt</i> (0.70/0.75)
Bord	Vit	Säng	Godis
Sitta	Natt	Dröm	Gullig
Dyna	Katt	Sova	Ful
Trä	Färg	Trött	Socket
Pall	Sammet	Vaken	Vacker
Kök	Dyster	Vila	Sur
Möbel	Sorg	Snarka	Snygg
Soffa	Hatt	Pigg	Kola
Bar	Hår	Slumra	Salt
Bekväm	Kol	Gäspa	Choklad
Bänk	Begravning	Dvala	Smak
Fåtölj	Död	Tupplur	Rar
Gunga	Bläck	Dåsig	Honung
Plats	Grått	Blunda	Syrlig
Säte	Sot	Ligga	Karamell
<i>Tjuv</i> (0.20/0.85)	<i>Tobak</i> (0.30/0.85)	<i>Tung</i> (0.50/0.75)	<i>Tyst</i> (0.30/0.55)
Polis	Röka	Lätt	Lugn
Inbrott	Cigarett	Bly	Högljudd
Bov	Pipa	Tjock	Stilla
Stjåla	Snus	Vikt	Bibliotek
Ta	Cancer	Kilo	Skrika
Kofot	Fest	Elefant	Mus
Fängelse	Marlboro	Fet	Viska
Olaglig	Filter	Hantel	Skogen
Sno	Tugga	Bak	Blyg
Stöld	Nikotin	Järn	Tomt
Bandit	Beroende	Massiv	Ro
Brottsling	Askfat	Bastant	Tiga
Kriminell	Fimp	Börda	Buller
Rån	Rulla	Belastning	Stum
Skurk	Bolma	Betong	Stiltje
<i>Vissla</i> (0.00/0.40)	<i>Whiskey</i> (0.30/0.80)	<i>Önska</i> (0.50/0.60)	<i>Örn</i> (0.20/0.35)
Ljud	Sprit	Jul	Fågel
Mun	Dryck	Vilja	Flyga
Melodi	Starkt	Få	Fri
Hund	Skottland	Present	Blick
Nynna	Flaska	Hoppas	Näbb
Tåg	Full	Brunn	Näste
Flöjt	Alkohol	Be	Falk
Blåsa	Malt	Födelsedag	Hök
Domare	Irländsk	Ande	Fjäder
Läppar	Pub	Fe	Jaga
Gnola	Rökig	Tanke	Uggla
Tuta	Bourbon	Trollspö	Uv
Nöjd	Torv	Radioprogram	Sväva
Kalla	Destilleri	Uppfyllelse	Sork
Tjuta	Konjak	Stjärnfall	Symbol