Original Papers

Polish Psychological Bulletin
2013, vol 44(4), 449-456
DOI - 10.2478/ppb-2013-0048

*Joanna Ulatowska**
*Justyna Olszewska***

Creating associative memory distortions - a Polish adaptation of the DRM paradigm

Abstract: *One of the most widely applied techniques used to examine associative memory errors is the Deese-Roediger-McDermott (DRM) paradigm. The aim of the present studies was to demonstrate a Polish version of the DRM paradigm and to test the characteristics of memory illusions evoked by this procedure for both recall and recognition. A normative study was conducted to prepare Polish stimuli material sharing similar characteristics as the lists in the English language version. Subsequently, the lists were applied to examine the effect of prior recall on recognition, as well as the influence of retention interval on recall. The results revealed that the Polish version of the DRM paradigm induced a robust effect of false recall and recognition. Moreover, it was revealed that immediate recall of a single list led to a higher rate of both correct and false recall and that prior recall positively influenced recognition, leading to a higher rate of hits.*

Key words: *the DRM procedure, false recognition, false recall, false memories*

Memory illusions, which have fascinated researchers for decades, refer to situations in which a person either declares that he or she remembers something that did not really occur or remembers a fact that did occur but in a manner that seriously differs from actually experienced events (Roediger, 1996). Studies have shown that memory errors arise from various psychological processes. Mazzoni (2002), for example, distinguished between *naturally occurring* and *suggestion-dependent memory distortions*. The first type of memory faults appears due to the natural processes of how memory works, whereas the second type occurs when there is an external source of suggestion. Nevertheless, these types of memory errors should not be treated as mutually exclusive categories because their boundaries are rather fuzzy and may overlap (Gobbo, 2002). Although memory distortions have been studied in the light of various paradigms (Mazzoni, 2002), one of the most widely applied is a laboratory technique developed to investigate associative memory errors, introduced by Roediger and McDermott (1995). This technique is a modification of the procedure previously utilized by Deese (1959) and is now known as the Deese-Roediger-McDermott (DRM) paradigm.

The aim of the present studies was to demonstrate a Polish version of the DRM paradigm and to test the characteristics of associative memory illusions evoked by the DRM procedure for both recall and recognition.

The objective of Deese's (1959) series of experiments was to test how associative factors affected recall. The participants were presented with lists of semantically associated words (e.g. thread, pin, eye, sewing, sharp, and point) and subsequently asked to recall them. The results revealed that when the associative bonds between list items were stronger, the participants were more prone to recall falsely the same common associate – the so-called critical lure (e.g. needle).

However, Deese's (1959) work did not gain recognition until the mid-1990s when Roediger and

* Institute of Applied Psychology, Academy of Special Education, ul. Szczęśliwicka 40, 02-353 Warszawa, julatowska@aps.edu.pl
** Institute of Applied Psychology, University of Social Sciences, ul. Sienkiewicza 9, 90-113 Łódź, justynao@umich.edu

This research was supported by The National Science Centre Grant DEC-2011/01/D/HS6/05482

McDermott (1995) replicated and extended his work and showed that the participants erroneously recalled the critical lures with similar probability as they recalled studied items from the middle of the lists. Moreover, the subsequent recognition test revealed a high rate of false alarms to non-presented critical lures and a very low false alarm rate to unrelated items. In addition, the phenomenology of false memories (Tulving, 1985) suggested that most subjects had experienced the critical lure equally vividly as the studied words. In a follow-up study, McDermott and Roediger (1998) revealed the robust nature of that illusion. They showed that warning the participants explicitly against associative errors did not eliminate false alarms to critical lures.

Associative memory errors, induced by the DRM paradigm, are explained in the light of various theoretical approaches. It is suggested that a critical lure is consciously or unconsciously (Underwood, 1965) activated during the encoding of its associates. The activation of a critical word leads to subsequent errors in recall and recognition that are caused by difficulties in identifying the source of the activation (Johnson, Hashtroudi, & Lindsay, 1993). According to the *source monitoring framework* (SMF; Johnson et al., 1993), the level of false memories should decrease when more distinctive characteristics are associated with studied items during encoding. This was supported by a number of studies that tested processes responsible for false recall and recognition reduction in the DRM paradigm. Small changes in the encoding phase, such as presentation modality (Gallo, McDermott, Percer, & Roediger, 2001; Smith & Hunt, 1998), slower presentation of lists (Gallo & Roediger, 2002; McDermott & Watson, 2001), or studying words along with their pictorial referents (Israel & Schacter, 1997), led to a substantial reduction in false memories rates. This suggests that more distinctive details enable participants to monitor differences between internally generated elements and those that derive from external sources (Johnson et al., 1993).

Since its introduction, the DRM paradigm has gained considerable attention, resulting in hundreds of studies. The popularity of the paradigm stems from the much simpler way of evoking false memories as compared to most other procedures which are usually complex both in terms of materials (e.g. videotapes) and time consuming (see Gerrie, Garry, & Loftus, 2005, for a review). Because of the compelling demonstration of memory distortions, the DRM paradigm is sometimes compared to a classic paradigm of interference production - the Stroop task (Gallo, 2010). Nowadays, Roediger and McDermott's (1995) work is one of the most often cited works in the field of memory distortions (cited 2205 times by November 2013 according to Google Scholar) and the DRM paradigm has been adopted in a broad range of domains, including the study of individual differences, aging, and neuropsychology.

The possibility of applying the DRM paradigm to a variety of domains of memory studies encouraged us to prepare the Polish version of this procedure. Because of possible cultural differences in word associations, a simple translation of the English language lists was not advisable (see Anastasi, De Leon, & Rhodes, 2005). Thus, the main goal of the present studies was to prepare Polish stimuli material with the same characteristics as the lists in the English language version. Our next aim was to test the newly created lists and their sensitivity to induce false recall and recognition.

Normative study

Method

Participants. A total of 215 undergraduate students (19–20 years of age) participated in the normative study. All were native Polish speakers.

Materials and procedure. Due to possible cultural differences, we could not use a Polish translation of the DRM lists (Deese, 1959; Roediger & McDermott, 1995) or of Russell and Jenkins's (1954) word association norms. To create the word lists that were later used in the main experiment the Kent-Rosanoff Word Association Test was applied. First, all 100 words from the test were translated into Polish by two independent translators. All discrepancies in the translation were discussed and one form of translation was chosen. Next, the participants were asked to write down their first association for each word. After that association the frequency ranking for each word was prepared. If different forms of words were given, the most frequent was selected and utilized.

Twelve lists were eliminated from further analyses as they did not reach the level of at least 15 different associations. Unfortunately there were many overlaps between the lists: that is, the same word appeared in two or more different lists. The only way to eliminate that constraint was to compose sets of lists in which none of words were overlapping. We were able to compose four different sets, made up of eight different lists. Each list consisted of 16 words: the stimulus word from the Kent-Rosanoff Word Association Test (which then served as a critical lure) and the first 15 associates of the critical word. The final four sets of lists are presented in the Appendix.

Main experiment

The main purpose of this experiment was to confirm that the lists created in the normative study would lead to a comparable level of false memories as in previous experiments that applied the DRM procedure (e.g. Roediger & McDermott, 1995) and to establish which set of lists would produce the highest intrusion rates. The second aim of the study was to examine how recall affects subsequent recognition. Thus, we examined recognition for lists that had previously been recalled as well as for lists that had not been recalled. Roediger and McDermott (1995) applied similar conditions so we followed this way of testing memory expecting comparable results: higher correct recognition rate for studied items and higher rate for critical lures in recognition after previous recall than in mere recognition condition.

Based on evidence suggesting that accurate recall decline faster over time than false recall (see Thapar & McDermott, 2001) our third purpose was to test recall performance after all study lists were encoded. We hypothesize that the recall rate for critical lures should remain stable whereas the rate of correct recall should visibly decrease when free recall is tested after encoding all study lists as compared to the condition in which free recall is provided after encoding each list separately. This led us to apply three different study-test conditions (delayed recall, recognition, recall + recognition) however, they let us receive four different types of data: for delayed recall (after encoding all study lists), for recognition (after encoding all study lists), for recognition after recall (recognition tested after each list are recalled), and for immediate recall (recall tested after encoded each list)[1] .

Method

Participants. A total of 246 undergraduate students took part in the experiment.

Materials. Four sets of semantically related word lists were utilized between-subjects during encoding. Each set consisted of eight 15-word lists. The first word in the list was the strongest associate of a target (critical) word. Each set of lists appeared in three different study conditions.

Design and Procedure. Three conditions were tested in a between-subjects design. The participants were randomly assigned to one of these conditions which differed in terms of the way memory was tested: 1. recognition after presentation of whole set of eight lists (recognition); 2. recall after presentation of whole set of eight lists (delayed recall); 3. recall after presentation of each list followed by recognition after whole set of eight lists (recall + recognition)[2] . The participants were told that the experiment was designed to test their memory and that they would be presented with eight word lists to memorize. In each condition participants were shown one set of eight lists. The words were presented in blocks, in the order of decreasing associative strength with a target word, using a PowerPoint presentation. The duration of display for each word was two seconds with a one second blank screen between the words.

After being presented with a single list, in delayed recall as well as in recognition condition, the participants viewed a completed mathematics equation for 30 seconds and decided whether or not it was solved correctly. After they were presented with all eight lists, the subjects were instructed to solve complex arithmetical problems for two minutes. Subsequently the participants in the delayed recall condition were asked to recall as many previously presented words as they could. The participants in the recognition condition received sheets with a recognition test and were instructed to indicate whether each item was old (had been seen earlier on one of the studied lists) or new (had not appeared on the study lists).

The recognition test was constructed similarly as in previous studies (e.g. Roediger & McDermott, 1995) and consisted of 48 words: 24 studied and 24 non-studied items. The studied words were drawn from the first, eighth and tenth position of each list. Three types of non-studied items were used: a) eight critical lures, b) 12 words drawn from the first, eighth and tenth position of four non-studied lists, and c) four critical lures from non-studied lists. The words on the recognition test sheet were arranged randomly.

In the recall + recognition condition, after being presented with a single list, participants were asked to solve a mathematical problem. The time for that task was 30 seconds and it was followed with a recall task, in which participants were instructed to write down all recalled items and were given two minutes to do so. After the last list was memorized and recalled, the participants were asked to solve arithmetical problems for two minutes. Next, the subjects received sheets with a recognition test identical to that used in the recognition condition. After completing the tasks, the participants in each condition were debriefed.

Results and discussion

The results will be presented separately for recall and recognition and for both types of test study conditions (i.e. recall + recognition vs. delayed recall and recall + recognition vs. recognition) will be compared. It lets us compare immediate recall (data are taken from recall + recognition condition) with delayed recall and recognition with recognition after recall.

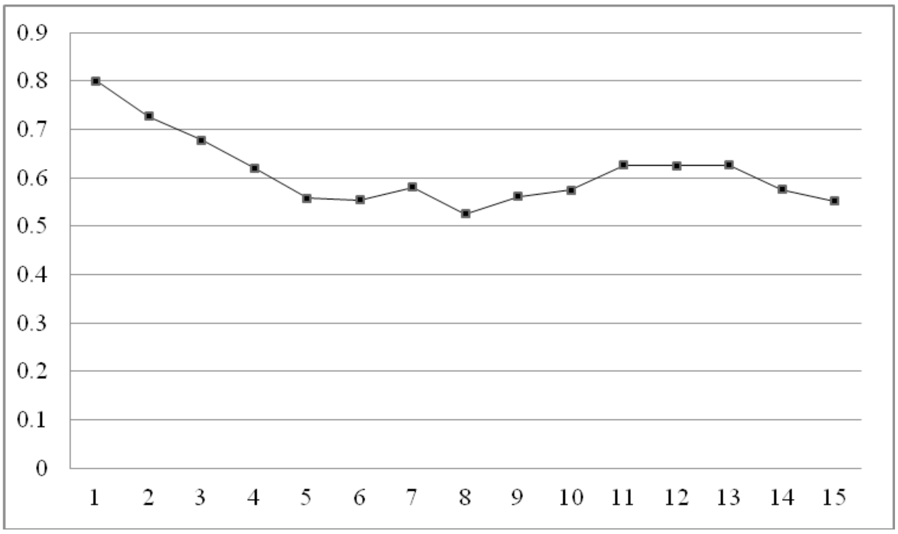
Recall data. The participants were able to recall studied items from all sets of lists with a mean rate of 0.64 (SD = 0.1) in immediate recall (data from the recall + recognition condition) and 0.22 (SD = 0.07) in the delayed recall condition. These rates differed significantly ($t(151) = 29.1$, $p < .001$, $d = 4.73$). This substantial difference between study conditions is not surprising as the participants in the delayed recall condition had to encode considerably more words at one time than the subjects in the recall + recognition condition (immediate recall), and the time between the beginning of encoding and the test was much longer for the former group.

For the recall + recognition condition (immediate recall) an analysis of the influence of serial position on correct recall was conducted. As in previous studies (Roediger & McDermott, 1995), a strong primacy effect was visible, which is the result of the order of words in the lists - the strongest associates of the critical words occurred early in the list (see Figure 1 - *page 452*). The recency effect was rather modest because, unlike in many previous studies, all participants had to solve the arithmetical task immediately after encoding each list and before free recall.

The critical lures in the recall + recognition condition were recalled with a probability of 0.29 (SD = 0.19), which is also a smaller proportion than in studies where immediate recall was provided (Johansson & Stenberg, 2002; Roediger & McDermott, 1995). The level of false recall was also much lower than the level of correct recall of the words from the middle of the lists (0.56 for positions 6–10). It should be stressed that our study involved visual presentation which may be the reason for discrepancies

[1] Recall + recognition condition gives us data for immediate recall and for recognition after recall. Therefore, while using the term immediate recall we refer to data from the recall + recognition condition.
[2] To remind: recall + recognition condition produces data for immediate recall and for recognition after recall.

Figure 1. Probability of correct recall as a function of serial position – all list sets.



in the results between our study and that of Roediger and McDermott (1995). However, the present data are in accord with Smith and Hunt (1998) who reported that studied items and critical lures were recalled at a mean rate of 0.72 and 0.22 respectively following a visual presentation.

Critical lures recall was lower in the delayed recall condition ($M = 0.20$; $SD = 0.15$) than in the recall + recognition condition (immediate recall) ($t(151) = 3.04$, $p < .01$, $d = 0.49$). This significant difference might again be the result of a discrepancy between conditions in the number of items studied simultaneously and the time between the beginning of encoding and the test. It can be seen that after a longer retention interval (delayed recall condition) false memories supersede true memories. This is consistent with Thapar and McDermott (2001) who showed that the decline in false recall and recognition was less pronounced over time than the decline in correct recall and recognition.

The probability of false recall of words other than critical lures was low in both conditions but significantly lower in the delayed recall condition ($M = 0.08$; $SD = 0.13$) than in the recall + recognition condition (immediate recall; $M = 0.14$; $SD = 0.17$; $t(151) = 2.49$, $p < .05$, $d = 0.40$).

For both conditions separately, an ANOVA was utilized to measure the differences between the lists' sets. In the recall + recognition condition (immediate recall) the mean proportion of accurate recall in set number three ($M = 0.72$; $SD = 0.08$) was significantly higher than in any other set ($F(3, 69) = 8.97$, $p < .001$, $\eta^2 = 0.28$). The proportion in the rest of the sets did not differ significantly from each other (set 1: 0.62, set 2: 0.56, set 4: 0.60). There were no significant differences between the sets of lists either in false recall of the critical lures ($F(3, 69) = .9$, n.s.) or in false recall of other words ($F(3, 69) = .78$, n.s.).

A similar pattern of results was obtained in the delayed recall condition. Once more the correct recall rate differed across conditions ($F(3, 76) = 12.29$, $p < .001$, $\eta^2 = 0.33$) and in set number 3 ($M = 0.26$; $SD = 0.06$) was significantly higher than in all other sets. Moreover, the difference in the proportion of correct recall was also significant between sets 1 ($M = 0.22$; $SD = 0.07$) and 2 ($M = 0.16$; $SD = 0.05$). The proportion of correctly recalled items in set 4 equalled 0.19 ($SD = 0.04$). No significant differences were observed between the sets of lists either in false recall of the critical lures ($F(3, 76) = .02$, n.s.) or in false recall of other words ($F(3, 76) = .54$, n.s.). The repeated pattern of results for all list sets in both conditions is evidence of the constant characteristics of sets. Although it is impossible to obtain sets of lists that have identical potential for critical word activation (see e.g. Roediger & McDermott, 1995; Roediger, Watson, McDermott, & Gallo, 2001; Stadler, Roediger, & McDermott, 1999), we can state that all of the presented sets are close to each other in evoking memory distortions and may be used as a part of the DRM paradigm.

Recognition data. The recognition test was conducted in both conditions - recognition and recall + recognition - after participants encoded all eight lists. The mean proportion of accurate recognition in the recognition + recall condition was high ($M = 0.79$; $SD = 0.14$) and comparable to other studies in which this condition was introduced. It was also significantly higher than in the recognition condition ($M = 0.63$; $SD = 0.19$; $t(164) = 5.75$, $p < .001$, $d = 0.90$). This is in accord with the hypothesis and previous studies (Roediger & McDermott, 1995, exp. 2) which proved that - also in a within-subjects design - the hit rate in the recall + recognition condition was greater than in the recognition condition. Such a result reflects the testing effect (Hogan & Kintsch, 1971; Karpicke, 2012; Roediger & Karpicke, 2006).

The false alarm rates to critical lures did not differ significantly in both conditions (recall + recognition: $M = 0.60$; $SD = 0.24$, and recognition: $M = 0.57$; $SD = 0.25$; $t(164) = 0.71$, n.s.). The lack of difference between the study conditions is not consistent with previous experiments (Roediger & McDremott, 1995) which revealed higher false

Table 1 Recognition results for all types of items as a function of list set

	Hits		False alarms			
			Critical lure		Non-related	
	R+R	R	R+R	R	R+R	R
Set 1	.76 (.18)a	.68 (.18)a	.51 (.27)a	.63 (.20)ab	.02 (.03)a	.01 (.04)a
Set 2	.79 (.07)ab	.68 (.12)a	.79 (.13)b	.74 (.18)a	.03 (.03)a	.02 (.04)a
Set 3	.86 (.10)b	.62 (.22)ab	.58 (.24)a	.50 (.28)bc	.02 (.05)a	.04 (.06)a
Set 4	.73 (.13)a	.54 (.19)b	.63 (.22)ab	.40 (.18)c	.06 (.06)b	.02 (.03)a

Note: R+R = recall + recognition condition, R = recognition condition; Standard deviations in parentheses; The rates with different indices differ significantly ($p < 0.05$).

alarm rate in the recall + recognition condition. However, this may be the result of the different modalities in which both studies were conducted.

The rate of false alarms to critical lures in both conditions (recall + recognition and recognition) in the present study was significantly lower than the hit rate (respectively $t(72) = 6.16$, $p < .001$, $d = 0.73$, and $t(92) = 2.86$, $p < .01$, $d = 0.28$). Although in Roediger and McDermott's (1995) experiment the rate of false alarms to critical lures was similar to the rate of hits, the present result is in accord with the outcome of Smith and Hunt's (1998) experiment 2 in which hits visibly exceeded critical lures, this being the result of different modality.

The rate of false alarms to non-studied items that were not associated with the studied lists was low in both conditions (recall + recognition: $M = 0.03$; $SD = 0.05$, and recognition: $M = 0.02$; $SD = 0.04$; $t(160) = 1.11$, n.s.).

Once more, the differences between the sets of lists were tested. The ANOVA conducted on both, the recall + recognition and the recognition condition showed significant differences between sets in the correct recognition of studied words ($F(3, 69) = 3.77$, $p < .05$, $\eta^2 = 0.14$, and $F(3, 89) = 2.85$, $p < .05$, $\eta^2 = 0.09$ respectively) and the false alarm rate for critical lures ($F(3, 69) = 3.28$, $p < .05$, $\eta^2 = 0.12$, and $F(3, 89) = 9.58$, $p < .001$, $\eta^2 = 0.24$ respectively). The rate of false alarms to foil items was significantly different only in the recall + recognition condition ($F(3, 68) = 3.58$, $p < .05$, $\eta^2 = 0.14$) but not in the recognition condition ($F(3, 86) = 1.35$, n.s.). All means and post hoc analysis results are displayed in Table 1.

General discussion

The present study was conducted to adjust the DRM paradigm to the Polish language and cultural conditions and to test the false recall and recognition effect it evokes. To summarize our findings present storng evidence that the Polish version of the DRM paradigm can elicit high levels of false recall and false recognition. Comparing our data to the DRM paradigm literature, we may point out slight differences in the rate of falsely recalled and recognized items. For instance, Roediger and McDermott's (1995) study revealed that, on an immediate test, critical lures were falsely recognized with the same probability as list words were correctly recognized. However, they used auditory presentation, whereas in our study we used visual presentation. Our data are in accord with those of Smith and Hunt (1998) who reported fewer falsely recalled items following visual presentation. A similar pattern of results pertains to recognition in that we noticed a slight decrease in the false recognition of critical lures compared to the original study (Roediger & McDermott, 1995), which again might be the result of using a visual presentation.

Our second interest was to test the effect of recall on recognition. Furthermore, we examined recall after short (immediate recall) and long retention intervals (delayed recall). The results were consistent with the hypotheses. First, it was revealed that immediate recall of a single list (i.e. consisting of 15 elements) led to higher rates of both correct and false recall. This supports parts of studies which report an increase in false memories (Thapar & McDermott, 2006). Our results illustrate well the phenomenon in which false memories supersede true memories with increasing retention intervals. This means that distinguishing between false and true memories becomes more difficult with time.

Second, prior recall positively influenced recognition, leading to a higher rate of hits. This reflects the testing effect (Hogan & Kintsch, 1971; Roediger & Karpicke, 2006) which says that memory testing enhances later retention. Prior recall, however, did not influence the false alarm rate, which might be a result of the visual modality used in this experiment. In Smith and Hunt's (1998) experiment 2, false recognition occurred at the rate of 0.45 in a similar condition to that in this study (recall + recognition). It is lower than in our study (0.60), but we assume that this discrepancy may stem from the different lengths of the lists used in these studies. We used 15 associated words in each list, whereas Smith and Hunt (1998) used 12. This might influence the activation of critical lures. Because Smith and Hunt (1998) did not introduce a recognition condition in their study, we do not have a point of reference for our result obtained in this condition. We may only assume that the modality of encoding is one of crucial factors in creating false memories.

As in the previous studies utilizing the DRM paradigm in languages other than English (e.g. Johansson & Stenberg, 2002), some differences in false recall and recognition rates to critical lures were observed as compared to the rate the original English lists produced. In our study[3], the rate of false recall was the lowest for lists associated with *fruit* (0%), *bath* (10%), and *lion* (10%). The highest false recall rates were generated for *joy* (65%), *window* (72%), and *sweet* (77%). The false alarm rates ranged from 31% for *soldier*, 38% for *mountain* and 39% for *lion* to 88% for *man*, 90% for *sweet* and 94% for *sickness*. Comparing the data obtained to the results of Stadler et al. (1999) we can see that part of the overlapping lists in the English original and Polish version led to a similar level of intrusion (e.g. *fruit* and *lion* which in both versions produced one of the lowest rates of false recall and *sweet* and *window* the highest) but most of them differed. Our results are in accord with the conclusion drawn by Johansson and Stenberg (2002) that the DRM lists should not simply be translated to other languages but rather created by means of the association test.

Taken together, our results confirm that applying the Polish version of the DRM paradigm led to both false recall and false recognition. Such material constitutes a well-controlled as well as utilitarian way of testing memory processes and affords the possibility of further investigating relationship between true and false memories in the DRM paradigm using native Polish speakers. We proposed and utilized four sets consisted of eight lists, however, other combinations are possible[4] depending on the study aim. Previous studies applied different numbers of lists, e.g., Roediger and McDermott (1995, experiment 1) used six lists and Gallo, McDermott, Percer and Roediger (2001, experiment 1) – 24 lists. Moreover, different number of items within a single list could be utilized (see Dodson & Hege, 2005; McCabe, Presmanes, Robertson, & Smith, 2004; Smith & Hunt, 1998). If application of smaller number of lists is possible, those producing higher levels of false recall and false recognition could be chosen (see Appendix).

Although the DRM paradigm is widely used in other countries (particularly in those where English is an official language) there is still much to discover in the associative memory distortions field. The DRM task, because of the easiness of evoking false memories, could be adopted in a variety of domains, including individual differences, neuropsychology or aging. More studies that investigate possible links between the DRM memory distortions and autobiographical memory phenomena (see Gallo, 2010) as well as susceptibility to other types of memory distortion are also needed (see Zhu, Chen, Loftus, Lin, & Dong, 2013). Moreover, providing the sets of Polish language lists allows cross-cultural comparisons and testing the influence of other potentially crucial variables e.g., bilingualism, on false memories. As Gallo (2010) has suggested the paradigm is also an example of a tendency in memory research to test the factors that rather influence retrieval quality not quantity.

References

Anastasi, J. S., De Leon, A., & Rhodes, M. G. (2005). Normative data for semantically associated Spanish word lists that create false memories. *Behavior Research Methods, Instruments, & Computers, 37*, 631-637.

Deese, J. (1959). On the prediction of occurrence of particular verbal intrusions in immediate recall. *Journal of Experimental Psychology, 58,* 17-22.

Dodson, C. S., & Hege, A. C. G. (2005). Speeded retrieval abolishes the false memory suppression effect: Evidence for the distinctiveness heuristic. *Psychonomic Bulletin & Review, 12*, 726-731.

Gallo, D. (2010). False memories and fantastic beliefs: 15 years of the DRM illusion. *Memory & Cognition, 38,* 833-848.

Gallo, D. A., McDermott, K. B., Percer, J. M., & Roediger, H. L. (2001). Modality effects in false recall and false recognition. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 27*, 339-353.

Gallo, D., & Roediger, H. L. (2002). Variability among word lists in eliciting memory illusions: Evidence for associative activation and monitoring. *Journal of Memory and Language, 4,* 469-497.

Gerrie, M., Garry, M., & Loftus, E. F. (2005). False memories. In: N. Brewer & K. Williams (red.), *Psychology and law. An empirical perspective.* New York: Guilford Press.

Gobbo, C. (2002). Occurring and Suggestion-Dependent Memory Distortions: Two separate research lines? *European Psychologist, 7*, 33-36.

Hogan, R. M., & Kintsch, W. (1971). Differential effects of study and test trials on long-term recognition and recall. *Journal of Verbal Learning and Verbal Behavior, 10,* 562–567.

Israel, L., & Schacter, D. L. (1997). Pictorial encoding reduces false recognition of semantic associates. *Psychonomic Bulletin and Review, 4*, 577-581.

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.

Johnson, M., Hastroudi, S., & Lindsay, D. (1993). Source monitoring. *Psychological Review, 114,* 3-28.

Karpicke, J. D. (2012). Retrieval-Based Learning: Active Retrieval Promotes Meaningful Learning. *Current Directions in Psychological Science, 21*, 157–163.

Mazzoni, G. (2002). Naturally Occurring and Suggestion-Dependent Memory Distortions: The Convergence of Disparate Research Traditions. *European Psychologist, 7*, 17-30.

McCabe, D. P., Presmanes, A. G., Robertson, C. L., & Smith, A. D. (2004). Item-specific processing reduces false memories. *Psychonomic Bulletin & Review, 11*, 1074-1079.

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.

McDermott, K. B., & Watson, J. M. (2001). The Rise and Fall of False Recall: The Impact of Presentation Duration. *Journal of Memory and Language, 45*, 160-176

Roediger, H. L. (1996). Memory illusions. *Journal of Memory and Language, 35*, 76-100.

Roediger, H. L., & Karpicke, J. D. (2006b). Test enhanced learning: Taking memory tests improves long-term retention. *Psychological Science, 17*, 249–255.

Roediger, H. L., & McDermott, K. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology, 21*, 803-814.

Roediger, H. L., Watson, J. M., McDermott, K. B., & Gallo, D. A. (2001). Factors that determine false recall: A multiple regression analysis.

[3] To enable comparison with other studies, these results are for the recall + recognition condition only.

[4] A verification of possible overlaps of words between the lists is advisable before the study.

Psychonomic Bulletin & Review, 8, 385-407.
Russel, 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, Office of Naval Research). Minneapolis: University of Minnesota.
Smith, R. E., & Hunt, R. R. (1998). Presentation modality affects false memory. *Psychonomic Bulletin & Review, 5*, 710-715.
Stadler, M. A., Roediger, H. L., & McDermott, K. B. (1999). Norms for word lists that create false memories. *Memory & Cognition, 27*, 494-500.
Thapar, A., & McDermott, K. B. (2001). False recall and false recognition induced by presentation of associated words: Effects of retention interval and level of processing. *Memory & Cognition, 29*, 424-432.
Tulving, E. (1985). Memory and consciousness. *Canadian Psychology, 26,* 1-12.
Underwood, B. J. (1965). False recognition produced by implicit verbal responses. *Journal of Experimental Psychology, 70*, 122-129.
Zhu, B., Chen, C., Loftus, E. F., Lin, C., & Dong, Q. (2013). The relationship between DRM and misinformation false memories. *Memory & Cognition*, published online 09 February 2013.

Appendix

32 lists used in the main experiment (arranged in 4 sets); rates of false recall and recognitions in parentheses (recall/recognition).

Set 1

Gwizdek *Whistle* (0.27/0.63)	**Kapusta** *Cabbage* (0.50/0.63)	**Kąpiel** *Bath* (0.1/0.45)	**Krótki** *Short* (0.18/0.50)	**Lekarz** *Doctor* (0.27/0.45)	**Okno** *Window* (0.72/0.68)	**Pień** *Stem* (0.13/0.4)	**Żołnierz** *Soldier* (0.13/0.31)
mecz	kiszona	wanna	długi	choroba	widok	drzewo	wojna
w-f	kwaszona	bąbelki	czas	fartuch	świat	las	mundur
sędzia	gołąbki	piana	mały	stetoskop	szyba	konar	wojsko
głośny	warzywo	relaks	odcinek	doktor	drzwi	mózg	karabin
dźwięk	bigos	woda	rękaw	kitel	firanka	kora	broń
trener	główka	czystość	sznurek	pomoc	przestrzeń	dąb	armia
hałas	dziecko	mydło	dystans	pielęgniarka	dom	drewno	odwaga
policjant	sałata	ciepła	dzień	szpital	duże	drwal	hełm
sport	groch	gorąca	ogon	słuchawki	rama	głuchy	służba
gwizd	obiad	przyjemność	patyk	recepta	zasłony	korzenie	Irak
świst	zielona	prysznic	włosy	specjalista	framuga	podstawa	przystojny
piłka	pusta	bańki	ołówek	zdrowie	niebo	siekiera	walka
sygnał	pole	basen	linijka	pediatra	wolność	śmierć	Afganistan
czajnik	głąb	morze	palec	pacjent	krajobraz	twardy	marionetka
koniec	jedzenie	odprężenie	sen	lek	otwarte	kłoda	militaria

Set 2

Igła *Needle* (0.4/0.8)	**Mężczyzna** *Man* (0.2/0.9)	**Motyl** *Butterfly* (0.2/0.6)	**Problem** *Problem* (0.4/0.9)	**Sen** *Sleep* (0.5/0.8)	**Słodki** *Sweet* (0.8/0.8)	**Sprawiedliwość** *Justice* (0.1/0.7)	**Stół** *Table* (0.3/0.8)
nitka	kobieta	owad	rozwiązanie	odpoczynek	cukierek	sąd	krzesło
szycie	seks	kolorowy	kłopot	łóżko	gorzki	prawo	obiad
ból	miłość	skrzydło	stres	noc	cukier	PiS	drewno
krew	człowiek	piękno	psycholog	marzenie	czekolada	waga	jedzenie
siano	przystojny	delikatny	trudność	przyjemność	lizak	dobro	obrus
strzykawka	siła	wiosna	zmartwienie	koszmar	smak	brak	nogi
ostra	bezpieczeństwo	lato	duży	jawa	ciastko	prawda	kuchnia
ukłucie	facet	ćma	smutek	spokój	kwaśny	uczciwość	blat
kompas	wysoki	kokon	trudny	relaks	słodycze	berło	dom
mała	mąż	larwa	zadanie	poduszka	miód	biały	łyżka
narkotyki	partner	latanie	hipoteza	błogość	pączek	cecha	posiłek
szpulka	chłopiec	poczwarka	konflikt	spać	baton	Hammurabi	mebel
świerk	tors	wolność	pomoc	mara	słony	honor	brąz
las	ojciec	barwny	walka	czuwanie	pyszny	idea	jadalnia
widły	spodnie	paź	wyzwanie	kamienny	smaczny	wartość	ława

Set 3

Głowa *Head* (0.13/0.47)	**Komfort** *Comfort* (0.27/0.78)	**Kwadrat** *Square* (0.17/0.6)	**Lew** *Lion* (0.1/0.39)	**Radość** *Joy* (0.65/0.65)	**Spragniony** *Thirsty* (0.34/0.6)	**Wysoki** *High* (0.39/0.63)	**Życzenie** *Wish* (0.3/0.52)
włosy	wygoda	koło	grzywa	uśmiech	woda	niski	marzenie
mózg	sklep	figura	król	szczęście	picie	mężczyzna	święta
ból	kanapa	matematyka	odwaga	smutek	pustynia	chłopak	urodziny
myślenie	fotel	trójkąt	zodiak	śmiech	sprite	brunet	prośba
mądrość	dywan	prostokąt	siła	miłość	napój	dąb	prezent
szyja	sofa	czerwony	ryk	zabawa	głodny	budynek	gwiazdka
pomysł	łóżko	geometria	Afryka	zadowolenie	sportowiec	koszykarz	kartka
kark	luksus	kształt	zwierzę	emocja	cola	poziom	spełnienie
pusta	relaks	pokój	tygrys	spokój	oaza	przystojny	magia
twarz	odpoczynek	blok	zoo	złość	potrzeba	siatkarz	rybka
kapelusz	meble	chi	kot	dzieciństwo	bieganie	szczupły	nadzieja
oczy	samochód	geometryczna	bajka	euforia	kac	wieżowiec	ostatnie
czapka	dyskomfort	kostka	dżungla	łzy	pepsi	długi	pragnienie
rozum	wypoczynek	romb	piękno	wolność	cytryna	słup	rozkaz
wiedza	poduszka	Rubik	sawanna	skakać	upał	drabina	wróżba

Set 4

Baranina *Mutton* (0.11/0.61)	**Choroba** *Sickness* (0.33/0.94)	**Głośny** *Loud* (0.27/0.66)	**Góra** *Mountain* (0.33/0.38)	**Księżyc** *Moon* (0.61/0.72)	**Kwitnąć** *Blossom* (0.11/0.77)	**Miękki** *Soft* (0.16/ 0.44)	**Owoc** *Fruit* (0/0.5)
mięso	ból	muzyka	dół	noc	kwiat	poduszka	jabłko
owca	rak	cichy	wysokość	pełnia	wiosna	twardy	banan
kebab	zdrowie	hałas	szczyt	gwiazdy	rozwijać	koc	słodki
jedzenie	łóżko	głośnik	śnieg	rogalik	rosnąć	puch	pomarańcza
baran	śmierć	dźwięk	Tatry	niebo	kobieta	miś	warzywo
danie	katar	krzyk	Zakopane	nów	rośliny	pluszowy	gruszka
jagnięcina	szpital	koncert	lód	słońce	dojrzewać	jedwab	truskawka
góral	lekarstwo	głos	Bieszczady	blask	ogród	miły	mandarynka
obiad	zło	dziecko	jezioro	piękno	róża	ciepły	soczysty
potrawa	lekarz	denerwujący	krajobraz	romantyczność	szczęście	gąbka	ananas
wełna	smutek	dyskoteka	morze	tajemniczość	ciąża	kot	cytryna
cielęcina	cierpienie	impreza	narty	czarodziejka	łąka	sweter	granat
talerz	grypa	klakson	przyroda	kratery	młodość	piórko	sok
wołowina	przewlekła	nieznośny	wspinaczka	spacer	więdnąć	przytulny	kiwi
wieprzowina	samopoczucie	rock	wycieczka	kosmos	wiśnia	aksamit	miąższ