

Retrieval dynamics in false recall: revelations from identifiability manipulations

Paula Carneiro · Angel Fernandez

Published online: 9 January 2013
© Psychonomic Society, Inc. 2013

Abstract The present study analyzed the retrieval dynamics of false recall, using an externalized free-recall task after participants studied Deese/Roediger–McDermott lists with high- and low-identifiable critical words. In Experiment 1, the memory test required participants to write down the words they remembered as having been presented in each list (recall output) plus any related words that came to mind (inclusion output). The results of the inclusion output showed that highly identifiable critical items were more frequently generated than less identifiable critical items, suggesting that highly identifiable critical words were more accessible in a first phase of retrieval. At the same time, the results of the recall output showed that highly identifiable critical items were less often falsely recalled than low-identifiable critical items, a replication of previous findings. In Experiment 2, self-reports corroborated that participants were using an editing strategy based on the identification and exclusion of critical words—that is, the identify-to-reject strategy. These results help us to more fully understand the identifiability effect and, beyond that, emphasize the importance of considering the intervening of dual processes of accessibility and error correction as a crucial feature in theoretical explanations of false memories.

Keywords False memory · DRM paradigm · Identifiability effect · Externalized free-recall · Self-reports

P. Carneiro
Faculty of Psychology, University of Lisbon,
Lisbon, Portugal

A. Fernandez
Faculty of Psychology, University of Salamanca,
Salamanca, Spain

P. Carneiro (✉)
Faculdade de Psicologia, Alameda da Universidade,
1649-013 Lisboa, Portugal
e-mail: mpcarneiro@fp.ul.pt

When lists of associated words (e.g., *door, glass, pane, shade, ledge, sill, house, open, curtain*, etc.) are studied without the converging word of each list (its critical word [CW]—e.g., *window*), participants in experiments often falsely remember that the nonstudied CWs were presented in the study lists (Deese, 1959; Roediger & McDermott, 1995). This simple procedure and this typical result characterize the Deese/Roediger–McDermott (DRM) paradigm. Although, in general, the results of the DRM paradigm are very impressive and robust, showing that false memories can be easily produced in laboratory settings, there is also evidence that false memories within this paradigm can be avoided. The latter is an interesting empirical finding, because it may further illuminate the mechanisms behind this type of memory errors both in the laboratory and in everyday life.

Some interesting findings, identifying conditions in which participants can reject false memories, have been reported. For example, some procedural manipulations lead to decrements in false recall or false recognition: the issuing of warnings at the time of study (McDermott & Roediger, 1998), slow presentation rate (McDermott & Watson, 2001), ample time to respond on the memory test (Benjamin, 2001), and, in general, manipulations that increase correct memory (e.g., repetition of lists; Brainerd, Reyna, Wright, & Mojardin, 2003). These manipulations are thought to facilitate error elimination because they increase the possibilities of using monitoring or verbatim memory. Moreover, some intrinsic characteristics of the CWs, such as length (e.g., Madigan & Neuse, 2004) and emotional value (e.g., Pesta, Murphy, & Sanders, 2001), can also make these words more distinctive than the studied words, contributing to false memory reduction.

More recently, another, more relational characteristic of the CWs has been demonstrated to be also important in false memory reduction: the identifiability of a given CW as a good referent for the theme connecting the words in its corresponding list (Carneiro, Fernandez, & Dias, 2009; Neuschatz, Benoit, & Payne, 2003). In work reporting this

identifiability effect, Carneiro et al. (2009, 2012) have shown that unrepresented CWs that are highly identifiable as the theme of their list are less likely to be falsely recalled and recognized. The explanation of the identifiability effect has been linked to the operation of the identify-to-reject (ITR) monitoring strategy (Gallo, 2006), which is based on assumptions about processes that take place during study and processes that take place at the time of the memory tests. Regarding encoding processes, it is assumed that CWs that are more identifiable are more likely to come to conscious attention while list words are studied and are tagged as not being part of the studied set. The main assumption regarding retrieval processes is that participants may adopt a monitoring strategy that edits out any item previously tagged as not presented for study.

A few other studies have already suggested that participants might use a strategy of this kind (Brédart, 2000; Jou & Foreman, 2007; Multhaup & Conner, 2002), but stronger converging evidence from direct manipulations of identifiability has been only recently available. For example, (1) a short presentation time for list words at study, a condition that is known to reduce conscious processing of the items, leads to the elimination of the identifiability effect (Carneiro et al., 2012, Experiment 1); (2) limited time at retrieval in a recognition test, a condition that is known to prevent participants from engaging in efficient monitoring, also gets rid of the identifiability effect (Carneiro et al., 2012, Experiment 2); and (3) young children, who are known to lack efficient monitoring abilities, do not show an identifiability effect (Carneiro et al., 2009).

However, the specific hypothesis that easily identifiable CWs are more likely to be accessible at the time of retrieval and, nonetheless, rejected via a monitoring process has not been directly tested in any of the previous studies. Performing such a direct test is the aim of the experiments reported here.

The hypothesis was tested using a procedure that had previously been employed to explore the dynamics of retrieval processes in recall, a procedure known as the *externalized free-recall task*. This task is based on the generate–recognize models of free recall (e.g., Anderson & Bower, 1972; Kintsch, 1970; Metcalfe & Murdock, 1981), which postulate that recall performance is determined by an initial generation stage and a final editing stage and has been used for understanding the dynamics of correct and error responses in recall tests (e.g., Unsworth, Brewer, & Spillers, 2010). It incorporates an inclusion or meaning instruction inspired by the studies of Jacoby (1991) and Brainerd and Reyna (1998a), in which all related words are allowed, minimizing the editing process and, thus, changing the recall output to an uninhibited recall (also used in the studies by Brainerd, Wright, Reyna, & Payne, 2002, and by Hege & Dodson, 2004). The procedure requires participants to output both words that they remember as

actually presented to them for study and related words that come to their mind during the test. They are also asked to differentiate the presented from the unrepresented words, either marking those that were actually presented (as in Hege & Dodson, 2004) or marking the unrepresented words that came to mind (as in Unsworth et al., 2010). This task permits the measurement of two types of output: the inclusion output, characterized by all the words generated, including studied and nonpresented related words, and the recall output, corresponding to the words that the participants recalled as studied words. Such a procedure permits registering items entering consciousness at the time of the test before they are lost in the final editing processes that filter the recall response. In this way, it is possible to tease apart two processes that occur at retrieval: item generation and item monitoring. In the two experiments reported next, this procedural logic was applied to the study of the mechanisms involved in the rejection of false memories.

Experiment 1

According to the ITR hypotheses, highly identifiable CWs could be more accessible and more likely to come to mind than low-identifiable CWs. As a result, in an experimental situation in which the memory test is an externalized free-recall task, highly identifiable CWs should be more likely to be produced in the generation phase of the test. Moreover, and corroborating previous results, it is expected that highly identifiable CWs will be less often produced as recall responses.

Method

Participants

Sixty-three students from Portuguese universities, majoring in psychology, economy, and hotel management, took part in the experiment. All were native Portuguese speakers (46 female; mean age = 21.3 years).

Design

The experiment followed a 2 (identifiability of the CW: high vs. low) \times 2 (type of response: recall vs. inclusion) factorial design, with repeated measures over the two factors.

Materials and procedure

Twelve associative Portuguese lists that produced extreme levels of CW identifiability in the study by Carneiro, Ramos, Costa, Garcia-Marques, and Albuquerque (2011) for backward associative lists were used. Theme identifiability was obtained in that normative study by asking

participants to generate, for each list, a single word that best defined its overall theme and was defined by the percentage of participants who produced the CW as the theme of the list. Six of these lists had CWs of high identifiability (CWs: *animal, money, fruit, heaven, eyes, and illness*), with percentages of theme identification ranging from 73.8 to 92.9. The other six lists had CWs of low identifiability (CWs: *high, slow, pen, dirty, anger, and warm*), with percentages of theme identification ranging from 1 to 11.2. The study lists consisted of the 10 strongest backward associates of each CW, arranged in decreasing order of strength. With the aim of controlling as much as possible for differential effects of lexical characteristics, the selected CWs of high- and low-identifiability lists did not differ in mean backward associative strength ($M_{\text{high}} = .21$ vs. $M_{\text{low}} = .19$), $t(10) = 0.68$, $p = .52$, frequency in the language ($M_{\text{high}} = 2,725$ vs. $M_{\text{low}} = 1,327$), $t(10) = 1.42$, $p = .19$, length ($M_{\text{high}} = 5.5$ vs. $M_{\text{low}} = 5.0$), $t(10) = 0.66$, $p = .53$, concreteness ($M_{\text{high}} = 5.5$ vs. $M_{\text{low}} = 4.7$), $t(10) = 1.5$, $p = .16$, or familiarity ($M_{\text{high}} = 5.7$ vs. $M_{\text{low}} = 5.2$), $t(10) = 1.5$, $p = .17$. Because of the small number of observations involved in each comparison, these statistical tests can be considered underpowered (the probability of finding a large statistical difference was .4). This fact does not by itself diminish the role that identifiability might play as the basis of a monitoring strategy, although it does not allow ruling out the possibility that identifiability could be related to one or more linguistic characteristics.

The participants were tested collectively in small groups. They were initially informed that they had to memorize lists of words for later recall. More specifically, they were instructed to write down the words they remembered as being presented in each list plus any related words that came to mind. To distinguish the presented from the related words, they were instructed to place a check mark only by the words that had been presented and to do it immediately after writing them. The 12 lists (6 of high and 6 of low identifiability) were randomly and auditorily presented to the participants at a rate of 1.5 s/word. After the presentation of each list, the participants had 1.5 min to write down, on a separate page of a 12-page booklet, the presented (with a check mark) and the related words for that list. The task lasted approximately 25 min.

Results

Table 1 displays the proportion of targets and CWs recalled for both types of response (inclusion and recall) and both types of lists (low and high identifiability). The recall response corresponds to the recalled items with a check mark, and the inclusion response corresponds to all the items produced in the task (all the items that came to mind). In general, high-identifiability lists showed higher generation of studied items (inclusion response) and also higher correct recall. For critical items, the

identifiability variable seemed to affect differently the inclusion and the recall outputs.

Two separate repeated measures ANOVAs were performed, one for studied items and another one for critical items. The analysis of studied items showed a main effect of type of response, $F(1, 62) = 66.00$, $MSE = .001$, $p < .001$, $\eta_p^2 = .52$, with the inclusion output producing higher levels of correct targets than the recall output ($M = .77$ vs. $M = .74$) and a main effect of identifiability, $F(1, 62) = 62.82$, $MSE = .006$, $p < .001$, $\eta_p^2 = .50$, with high-identifiable lists producing higher levels of correct targets than low-identifiable lists ($M = .79$ vs. $M = .72$). No interaction effect was found for studied items.

The analysis of critical items showed a main effect of type of response, $F(1, 62) = 118.43$, $MSE = .038$, $p < .001$, $\eta_p^2 = .66$, with the inclusion output producing more CWs than the recall output ($M = .37$ vs. $M = .11$). There was also a significant type of response \times identifiability interaction, $F(1, 62) = 18.13$, $MSE = .023$, $p < .001$, $\eta_p^2 = .23$, showing that the lists of high identifiability produced lower levels of false recall than did lists of low identifiability ($M = .07$ vs. $M = .15$), $t(62) = 3.48$, but generated (inclusion output) a higher number of CWs than did lists of low identifiability ($M = .42$ vs. $M = .33$), $t(62) = 2.08$. The post hoc analyses (Bonferroni tests) confirmed that the difference in the recall output was significant ($p < .001$), as well as the difference in the inclusion output ($p < .05$).¹

For completeness, an additional analysis was performed for the other words (neither studied nor critical) produced. These words corresponded to any response that could not be classified as studied or critical, and they were almost always related to the themes of the specific lists. Since these words were almost all gist-related words, it seemed appropriate to investigate whether the pattern observed for these words was different from the pattern found for critical items. Because the critical items might be the best themes of the lists, there is no reason to find differences in identifiability

¹ Two of the lists of high identifiability were lists in which their CWs were names of categories. Those were the cases of the *fruit* and *animal* lists, containing, respectively, six and seven exemplars belonging to each category. It is not very surprising that this type of list was of high identifiability, since in the presence of the exemplars, it is very easy to extract the name of the category as the theme of the list. Since those two lists are atypical and could lead to alternative interpretations, we have performed the same analysis excluding them. The same main effects of task and identifiability emerged for studied items, $F(1, 61) = 62.45$, $MSE = .006$, $p < .001$, $\eta_p^2 = .51$, and $F(1, 61) = 37.11$, $MSE = .006$, $p < .001$, $\eta_p^2 = .38$, respectively, as well as the same main effect of task for critical items, $F(1, 62) = 112.25$, $MSE = .057$, $p < .001$, $\eta_p^2 = .64$. More important, the same significant task \times identifiability interaction occurred for critical items, $F(1, 62) = 20.14$, $MSE = .057$, $p < .001$, $\eta_p^2 = .25$, with the post hoc tests showing that in the recall task, the lists of high identifiability produced lower levels of false recall than did lists of low identifiability ($M = .10$ vs. $M = .15$, $p < .05$), whereas in the inclusion task, the lists of high identifiability produced a higher number of CWs than did lists of low identifiability ($M = .45$ vs. $M = .33$, $p < .01$).

Table 1 Proportions of targets and critical words as a function of type of output (inclusion and recall) and type of lists (low and high identifiability) in Experiment 1 and Experiment 2

		Inclusion				Recall				
		High Ident.		Low Ident.		High Ident.		Low Ident.		
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	
Experiment 1	Targets	.80	.08	.73	.09	.78	.09	.70	.10	
	Critical	.42	.34	.33	.20	.07	.12	.15	.15	
Experiment 2	Targets	.80	.07	.75	.07	.78	.08	.71	.08	
	Critical	Total sample (<i>N</i> = 60)	.59	.30	.35	.20	.11	.14	.15	.15
		Identifiers (<i>n</i> = 25)	.71	.27	.37	.18	.07	.11	.17	.18
		Nonidentifiers (<i>n</i> = 35)	.51	.30	.34	.22	.13	.16	.14	.13

for other related words. The total of other words produced was divided by the number of lists. The number of these words that were incorrectly recalled as actually studied was low, not exceeding 2 per list, but the number of these words that were generated was relatively high, 14 per list. The same type of 2 (identifiability of CW: high vs. low) \times 2 (task: inclusion vs. recall) ANOVA showed a significant effect only of type of response, $F(1, 62) = 46.82$, $MSE = 1.75$, $p < .001$, $\eta_p^2 = .43$, indicating that, as was expected, in the inclusion output, more gist-related words were produced than in the recall output ($M = 1.36$ vs. $M = .22$).

Experiment 2

The pattern of results observed in Experiment 1 is clearly consistent with the ITR explanation of false memory reduction for identifiable CWs. Still, a stronger case for the use of such a strategy could be made if new, convergent evidence could be obtained. One reasonable approach, given the conscious, purposeful, and effortful nature of the hypothesized strategy, could be to ask of participants whether they have used it at all when performing the externalized recall task. And that is precisely what was done in the present experiment, in which participants went through the same study and test procedures as in the previous experiment and, additionally, responded to end-of-experiment questionnaires on the type of strategies they had used to maximize true memory and reduce false memory.

Method

Participants

Sixty psychology students from the University of Lisbon took part in this experiment. All were native Portuguese speakers (52 female; mean age = 21 years).

Materials and procedure

This experiment employed the same materials and procedure as those used in Experiment 1 but added a posttesting phase in which the participants were asked to report the strategies used during the task. Approximately half the participants ($n = 31$) were asked to freely report the strategies that they had used during the task in response to an open request: “please report any strategy that you have used to perform this task.” The other half of the sample ($n = 29$) were asked to respond to a longer questionnaire, in which they had to select the strategies used, marking them among several provided alternatives that were inspired in the self-reported strategies given by participants to the open question. This longer questionnaire included two main requests: “please select the strategies used to better memorize the words of the lists” and “please select the strategies used to distinguish the presented from the thought words.” The rationale for using these two questions was to analyze whether the identification of the theme of the list was used only to help the encoding of the list items or was also used to reject the themes (critical items) that were identified.

Results

The results of this experiment replicated the pattern of findings observed in Experiment 1 in regard to the generation and recall of studied and critical items (see Table 1). The two more important results found in the previous experiment were obtained in this one. First, in relation to targets, high-identifiable lists produced higher levels of correct inclusion/recall than did low-identifiable lists ($M = .79$ vs. $M = .73$), $F(1, 59) = 64.91$, $MSE = .003$, $p < .001$, $\eta_p^2 = .52$. Second, in relation to critical items, a significant type of response \times identifiability interaction, $F(1, 59) = 49.42$, $MSE = .024$, $p < .001$, $\eta_p^2 = .46$, was found. Lists of high identifiability generated higher levels of CWs than did lists of low identifiability

($M = .59$ vs. $M = .35$), $t(59) = 6.01$, but produced lower levels of false recall ($M = .11$ vs. $M = .15$), $t(59) = 2.12$, both comparisons significant under Bonferroni conditions.

The percentages of strategies reported in the open question and selected in the questionnaire are displayed in Table 2. When the participants were instructed to freely report the strategies that they used during the task, 29 % of them answered that they had identified the themes of lists. This result confirmed a similar result obtained by Gallo, Roberts, and Seamon (1997) showing that about 20 % of their participants in a no-warning condition reported on a posttest questionnaire that they had used a strategy aimed at determining the critical lures. This strategy was the second most frequent strategy referred to by the participants, second

to mentioning the order in which they had recalled the words (usually starting with the last words and then recalling the first words).

The responses to the questionnaire showed that the strategy of identifying the themes of the lists was highly selected as a strategy for memorizing the words of the lists (72 %), but also as an editing strategy (55 %). This editing strategy was the second most frequently selected strategy by the participants. Only one other editing strategy, based on memory source (remembering the absence of thoughts or images for the items) was reported with higher frequency (66 %).

In order to better understand the relationship between the reported theme identification strategies and the rates of generation/rejection of CWs, further statistical analyses were conducted. The data from participants who freely reported the identification of a theme in response to the open request ($n = 9$) and those who selected theme identification as an editing strategy ($n = 16$) were pooled into a single (identifiers) set, and the data from the rest of participants ($n = 35$) formed a contrasting (nonidentifiers) set (see Table 1). The inclusion output and the recall output for these two groups of participants were compared, separately for high- and low-identifiability CWs, with analyses showing that identifiers had higher levels of inclusion of highly identifiable CWs than did nonidentifiers ($M = .71$ vs. $M = .51$), $t(58) = 2.66$, $p < .01$, as well as a tendency to recall lower levels of highly identifiable CWs ($M = .07$ vs. $M = .13$), $t(58) = 1.73$, $p = .09$. The same comparisons performed for low-identifiable CWs between the two groups of participants showed no significant differences. A second analysis was conducted on the basis of the assumption that a good estimate of the use of the ITR strategy would be one reflecting the amount of cases in which the participant, first, generates the CW and, second, avoids marking it as a studied word (i.e., an inclusion-minus-recall score). The analysis of this measure for highly identifiable critical items showed that the group of identifiers scored much higher than the group of nonidentifiers ($M = .63$ vs. $M = .37$), $t(58) = 3.54$, $p < .01$. Finally, in a regression analysis aimed at assessing the predictive potential of the three main strategies reportedly used to distinguish between presented and unpresented words (recalling thoughts or images, identifying and rejecting theme words, and recalling presentation order), only the ITR strategy ($\beta = .539$) was a significant predictor of inclusion-minus-recall scores, $R^2 = .21$; $F(3, 25) = 3.50$, $p = .03$. Even taking into account the interpretive limitations of this type of analysis when small samples are used, when the results of the various ways of analyzing the data are considered, there is a strong indication that participants reporting strategies based on theme identification and exclusion were, in general, more successful in identifying and rejecting the CWs.

Table 2 Percentage of strategies reported

Open question	%
Recall the last and the first words (usually the last words first)	45
Identify the theme	29
Rehearsal	23
Create images	19
Create sentences	16
Create stories	16
Association to like/dislike	6
Think of opposite words	6
Recall the words in the presented order	3
Count the number of words	3
Questionnaire	
Strategies used to memorize the words of the lists	%
Rehearsal while listening	90
Recall first the last words and then the first words	76
Identify the theme	72
Create images	69
Recall the words in the presented order	45
Create sentences	28
Association to like/dislike	28
Count the number of words	21
Create stories	17
Memorize the sound	17
Memorize the initials of each word	3
No strategy	0
Other strategy	0
Strategies used to distinguish the presented from the unpresented words	%
Recalling thoughts or images	66
Identifying and rejecting theme words	55
Recalling presentation order	38
Generating synonyms or antonyms	21
Exhaustive recall	10
No strategy	7
Other	3

General discussion

The two experiments reported in this study provide new and direct evidence regarding the retrieval dynamics underlying identifiability-based rejection of false memories, leading to identification of the mechanisms that produce the identifiability effect and, more important, to a better understanding of how false memories are produced and avoided in the DRM paradigm. The findings of Experiment 1 show that critical items that are easily identified as the themes of the lists, although less likely to be produced in final recall, have a marked tendency to come to mind at the time of retrieval. This finding is evident in the externalized memory task used in Experiment 1 and also in the verbal reports of a significant number of participants undergoing the same task in Experiment 2. Although dynamic interplay between generation and elimination operations has been earlier assumed to be at the foundation of strategic reduction of false memories for highly identifiable lures (see Carneiro et al., 2009, 2012), this is the first study to provide direct evidence, stemming from both memory tests and confirming self-reports, that the identify-to-reject is very often the strategy of choice.

The pattern of generation/recall for critical items was found to be very different from the pattern of generation/recall for related items, which suggests that the results are in fact due to the manipulation of theme identifiability restricted to the critical items. Whereas we found an interaction effect between type of response and identifiability for critical items, the same did not occur for other related items. However, something must be said about the low levels of false recall obtained in this experiment ($M = .11$). One of the reasons that could explain it could be linked to the fact that lists were presented with only 10 associates, and it is known that false recall decreases as the list length reduces (Robinson & Roediger, 1997). Moreover, we selected the higher and lower lists in identifiability that had similar values of BAS, and, due to the almost nonexistence of low identifiable lists with high BAS, we were forced to select all lists with relatively low BAS ($M = 1.9$). Another reason might be related to the procedure that was used, which in some way could have worked as a warning. Thus, the three factors—lists with fewer associates, relatively low BAS, and the specific procedure—could explain why we obtained low levels of false recall. However, for the purpose of the present study, this does not seem to be a limitation.

It is reasonable to conclude that the specific experimental task used in this study, demanding inclusion plus recall responses, corresponds to what internally happens when participants are studying DRM lists and instructed to recall each list. Usually, in a standard recall task, we only have access to the retrieval output already filtered by some sort of memory editing, and thus it is difficult to measure the contribution of the different processes involved in false

memory formation. But with the methodology of the present study, we have access to the generation or stimulation phase, corresponding to all the words that come to mind at retrieval (inclusion output), and then to the retrieval measure already free from the items that were rejected by monitoring processes (recall output). It should be pointed out that although the stimulation phase is usually considered to occur at encoding and error editing at retrieval, the experimental procedure used in the present experiment allows for both mechanisms to be operating at retrieval. Accessibility of CWs at retrieval can be due to a reactivation of the CW generated at encoding or can be caused by a CW generation at test. Although the present study does not provide a definitive answer regarding this issue, the fact that the CWs were mainly positioned on the last quintile in the inclusion and recall outputs (59 % and 31 % of the time, respectively) suggests that at least some CWs could have been generated at test (as first proposed by Roediger & McDermott, 1995). In further research, it would be very interesting to study the contributions of CWs' accessibility at both stages. In that regard, using think-out-loud procedures (see Lampinen, Meier, Arnal, & Leding, 2005; Seamon et al., 2002) could be procedural additions of high value.

On the basis of previous findings and the evidence reported in the present study, the pattern of results needs to be explained by a dual-process approach that incorporates a process that stimulates the CWs and another process that eliminates them (e.g., Brainerd & Reyna, 1998b). For example, if one accepts that gist could be operationally specified by the extraction of the list's theme, these results can be easily explained by the fuzzy-trace theory. For fuzzy-trace theory, the extraction of meaning (gist) from the list items promotes error inflation, but a recollection rejection process could suppress false memories by making use of available verbatim traces (Brainerd et al., 2003). We have shown here that lists with a stronger thematic representation or gist would lead to higher generation of the critical items. In addition, we found that these lists also produced more correct recall, which could explain why they also lead to higher rejection of false memories: On the basis of the verbatim trace of studied items, the process of recollection rejection could have suppressed false memories for highly identifiable lists (Brainerd et al., 2003). An alternative theoretical stand such as the activation/monitoring framework (Roediger, Balota, & Watson, 2001), at its base also a dual-process view, could possibly accommodate the results. Still, the fact that the identifiability effect is found even when backward associative strength is controlled across type of lists demands additional theoretical elaboration, incorporating the possibility of overadditivity effects modulating and raising the mean BAS of lists formed by a set of items strongly related to a unique theme word (cf. Watson,

Balota, & Roediger, 2003), considering the effects of other types of associative variables (e.g., density of interconnections among studied items, as in McEvoy, Nelson, & Komatsu, 1999), or admitting the possibility that item activations could derive from mechanisms other than associative-based automatic propagation.

In summary, studying the dynamics of retrieval helps us to more fully understand the identifiability effect by examining the processes that stand behind it, and it contributes to a deeper comprehension of the processes underlying memory distortion phenomena. The results strongly reinforce theoretical approaches that explain false memories by the intervention of two processes, emphasizing, as suggested by Gallo (2010), the need to adopt a multiple-process view in this field of research while, at the same time, using converging methods that may result in more exact characterization of the contributing components.

Author Note The work of both authors was supported by the Portuguese Foundation of Science and Technology (Project PTDC/PSI-PCO/112466/2009). The work of Angel Fernandez was also supported by the Spanish Ministry of Science and Innovation (Project PSI2008-05607/PSIC). The suggestions by the journal action editor (I. Dobbins) and by anonymous reviewers of an earlier version of the manuscript are greatly appreciated.

References

- Anderson, J. R., & Bower, G. H. (1972). Recognition and retrieval processes in free recall. *Psychological Review*, *79*, 97–123.
- Benjamin, A. S. (2001). On the dual effects of repetition on false recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*, 941–947.
- Brainerd, C. J., & Reyna, V. F. (1998a). Fuzzy-trace theory and children's false memories. *Journal of Experimental Child Psychology*, *71*, 81–129.
- Brainerd, C. J., & Reyna, V. F. (1998b). When things that never happened are easier to “remember” than things that did. *Psychological Science*, *9*, 484–489.
- Brainerd, C. J., Reyna, V. F., Wright, R., & Mojardin, A. H. (2003). Recollection rejection: False memory editing in children and adults. *Psychological Review*, *110*, 762–784.
- Brainerd, C. J., Wright, R., Reyna, V. F., & Payne, D. G. (2002). Dual-retrieval processes in recall. *Journal of Memory and Language*, *46*, 120–152.
- Brédart, S. (2000). When false memories do not occur: Not thinking of the lure or remembering that it was not heard? *Memory*, *8*, 123–128.
- Carneiro, P., Fernandez, A., & Dias, A. R. (2009). The influence of theme identifiability on false memories: Evidence for age-dependent opposite effects. *Memory & Cognition*, *37*, 115–129.
- Carneiro, P., Fernandez, A., Diez, E., Garcia-Marques, L., Ramos, T., & Ferreira, M. B. (2012). “Identify-to-reject”: A specific strategy to avoid false memories in the DRM paradigm. *Memory & Cognition*, *40*, 252–265.
- Carneiro, P., Ramos, T., Costa, R. C., Garcia-Marques, L., & Albuquerque, P. (2011). Identificabilidade dos temas de listas associativas *backward*: Contributos para o estudo das memórias falsas. *Laboratório de Psicologia*, *9*, 23–34.
- Deese, J. (1959). On the prediction of occurrence of particular verbal intrusions in immediate recall. *Journal of Experimental Psychology*, *58*, 17–22.
- Gallo, D. A. (2006). *Associative illusions of memory: False memory research in DRM and related tasks*. New York: Psychology Press.
- Gallo, D. A. (2010). False memories and fantastic beliefs: 15 years of the DRM illusion. *Memory & Cognition*, *38*, 833–848.
- Gallo, D. A., Roberts, M. J., & Seamon, J. G. (1997). Remembering words not presented in lists: Can we avoid creating false memories? *Psychonomic Bulletin & Review*, *4*, 271–276.
- Hege, A. C. G., & Dodson, C. S. (2004). Why distinctive information reduces false memories: Evidence for both impoverished encoding and distinctiveness heuristic accounts. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*, 787–795.
- Jacoby, L. L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory and Language*, *30*, 513–541.
- Jou, J., & Foreman, J. (2007). Transfer of learning in avoiding false memory: The roles of warning, immediate feedback, and incentive. *The Quarterly Journal of Experimental Psychology*, *60*, 877–896.
- Kintsch, W. (1970). Models for free recall and recognition. In D. A. Norman (Ed.), *Models of human memory* (pp. 307–373). New York: Academic Press.
- Lampinen, J. M., Meier, C. R., Arnal, J. D., & Leding, J. K. (2005). Compelling untruths: Content borrowing and vivid false memories. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*, 954–963.
- Madigan, S., & Neuse, J. (2004). False recognition and word length: A reanalysis of Roediger, Watson, McDermott and Gallo (2001) and some new data. *Psychonomic Bulletin & Review*, *11*, 567–573.
- 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.
- McEvoy, C. L., Nelson, D. L., & Komatsu, T. (1999). What is the connection between true and false memories? the differential roles of interitem associations in recall and recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*, 1177–1194.
- Metcalfe, J., & Murdock, B. B. (1981). An encoding and retrieval model of single-trial free recall. *Journal of Verbal Learning & Verbal Behavior*, *20*, 161–189.
- Multhaup, K. S., & Conner, C. A. (2002). The effects of considering nonlist sources on the Deese-Roediger-McDermott memory illusion. *Journal of Memory and Language*, *47*, 214–228.
- Neuschatz, J. S., Benoit, G. E., & Payne, D. G. (2003). Effective warnings in the Deese-Roediger-McDermott false-memory paradigm: The role of identifiability. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*, 35–41.
- Pesta, B. J., Murphy, M. D., & Sanders, R. E. (2001). Are emotionally charged lures immune to false memory? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*, 328–338.
- Robinson, K. J., & Roediger, H. L. (1997). Associative processes in false recall and recognition. *Psychological Science*, *8*, 231–237.
- Roediger, H. L., Balota, D. H., & Watson, J. M. (2001). Spreading activation and the arousal of false memories. In H. L. Roediger, J. S. Nairne, I. Neath, & A. M. Surprenant (Eds.), *The nature of remembering: Essays in honor of Robert G. Crowder* (pp. 95–115). Washington: American Psychological Association.
- 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.

- Seamon, J. G., Lee, I. A., Toner, S. K., Wheeler, R. H., Goodkind, M. S., & Birch, A. D. (2002). Thinking of critical words during study is unnecessary for false memory in the DRM procedure. *Psychological Science, 13*, 526–531.
- Unsworth, N., Brewer, G. A., & Spillers, G. J. (2010). Understanding the dynamics of correct and error responses in free recall: Evidence from externalized free recall. *Memory & Cognition, 38*, 419–430.
- Watson, J. M., Balota, D. A., & Roediger, H. L. (2003). Creating false memories with hybrid lists of semantic and phonological associates: Over-additive false memories produced by converging associative networks. *Journal of Memory and Language, 49*, 95–118.