MEMORY, 2017
VOL. 25, NO. 8, 986–998
http://dx.doi.org/10.1080/09658211.2016.1239742

Routledge
Taylor & Francis Group

REVIEW ARTICLE

Check for updates

Explaining the persistence of false memories: a proposal based on associative activation and thematic extraction

Paula Carneiro[a], Leonel Garcia-Marques[a], Ana Lapa[a] and Angel Fernandez[b]

[a]Faculty of Psychology, University of Lisbon, Lisbon, Portugal; [b]INICO, University of Salamanca, Spain

ABSTRACT

This study aimed to analyse the effect of retention intervals on associative and thematic false memories. Two experiments, using two types of critical items that were either associatively or thematically related to studied material, were conducted. In both experiments, one group of participants performed a recognition test immediately after the presentation of lists, and another group performed the task one week later. In Experiment 1, the recognition test consisted of pairs of items with four response alternatives (*both items had been presented, only the left item had been presented, only the right item had been presented* or *none of the items had been presented*). Critical items were also manipulated so that they were either presented in or absent from the list. In Experiment 2, a standard recognition test that differed in the mode of presentation was used: self-paced or speeded response. Both experiments showed that associative critical items were more recognised than thematic critical items in the immediate condition. However, whereas associative critical items decayed after a one-week delay, thematic critical items were similarly recognised at both retention intervals. The findings of the present study suggest that each type of process – associative and thematic – behave differently over time.

ARTICLE HISTORY
Received 13 January 2016
Accepted 16 September 2016

KEYWORDS
False memories; DRM paradigm; retention interval; associative activation; thematic extraction

False memories have been studied by using controlled laboratory paradigms, such as the Deese–Roediger–McDermott (DRM) paradigm (Deese, 1959; Roediger & McDermott, 1995). In this paradigm, participants are instructed to listen to lists of words, each of them consisting of strong associates of a nonpresented word, the critical item (CI), for a later memory test. The results of numerous memory studies using this procedure have shown participants often recall and recognise the critical nonpresented words as having been presented in the study phase (Gallo, 2006, 2010). A surprising finding in the false memory literature is that false memories are persistent and usually more stable than true memories (for a review of studies, see Gallo, 2006). However, the processes that promote this longer persistence are not entirely clear. In this study, we address this issue by analysing the effect of a delay on items that are not presented but are representative of either associative activation or thematic extraction of the presented material.

Many studies have analysed the effect of test delays (hours, days, weeks or even months) by using recall and recognition DRM tasks (e.g., McDermott, 1996; Seamon, Luo, Kopecky, et al., 2002; Thapar & McDermott, 2001; Toglia, Neuschatz, & Goodwin, 1999). The results have been mixed but, almost all these studies have shown a reduction in the levels of correct recall and correct recognition with longer retention intervals and, interestingly, that false memories are less affected by a time delay than true memories. Despite this general pattern, the results have been more consistent with recall tests than with recognition tests (Gallo, 2006).

Despite this general pattern of results we cannot dismissed the results of other studies that found a more surprising pattern, with false memories not affected at all (Toglia et al., 1999) or with an increase in falsely recognised pictures of category exemplars after a 3-day delay (Seamon, Luo, Schlegel, Greene, & Goldenberg, 2000).

These findings have been important in testing theories of false memories. There are mainly two theories that explain the DRM phenomenon. One is the activation-monitoring framework (Roediger, Balota, & Watson, 2001), mainly inspired by spreading activation models (e.g., Collins & Loftus, 1975) and by the concept of source monitoring (Johnson, Hashtroudi, & Lindsay, 1993). Briefly described, this framework adopts the idea that the processing of one word activates a corresponding node in our mental lexicon and that this activation spreads to surrounding concept nodes. In the case of the DRM illusion, when studied words are processed, a summation of activation occurs towards the CI, thus frequently producing a false memory. However, another part of the process – source monitoring – has also its contribution. According

CONTACT Paula Carneiro mpcarneiro@fp.ul.pt, mpcarneiro@hotmail.com Faculdade de Psicologia, Alameda da Universidade, 1649-013 Lisboa, Portugal

to the source-monitoring framework, a memory is encoded with its specific attributes, including perceptual attributes (i.e., sensory, temporal, and spatial information) and the cognitive operations occurring at the moment of encoding. When the source of the error-inflated memory is confused with the source of true memories, a false memory can occur (Johnson et al., 1993). Only when the sources of the true and the error-inflated memories are distinguished and one can attribute the origin of a false memory to an internal generation process instead of to an actual external world event, a false memory can be rejected. The effect of an efficient source monitoring is greatly exemplified in studies in which the participants were warned about the DRM effect before study (e.g., McDermott & Roediger, 1998). In general, the studies showed that warned participants can successfully reject a great amount of false memories.

Fuzzy trace theory (Brainerd & Reyna, 2005; Reyna & Brainerd, 1995) is the other most prominent theory, and it states that people encode two different traces of information: a gist trace, which captures the basic meaning of the event, and a verbatim trace, which preserves the specific surface form of information. This theory attributes the production of false memories to gist processing, which in the case of the DRM illusion, would be the extraction of the connected meaning of the studied words regarding the lists' theme. On the other hand, verbatim processing has the function of suppressing or rejecting false positives.

These findings indicating that false memories tend to be less affected by a time delay than true memories suggest that theories solely based on activation cannot easily explain the consistently observed false-persistence effects, as it is well known that automatic activation rapidly decays with time (Anderson, 1983). Indeed, lexical decision studies have shown that associative priming is transitory and dissipates rapidly (Neely, 1977). Tse and Neely (2005) used items in DRM lists as primes for CI targets and showed that this type of list, since they comprise many associates, produce more stable priming than single associates of CI targets. However, priming in this study lasted for less than 1 minute, which could not explain the effect of the retention interval of hours, days or weeks with respect to false memories, as has been found in retention time studies. Moreover, a study by Cotel, Gallo, and Seamon (2008) corroborated this finding, showing that priming for CIs can be observed immediately after a single DRM list was studied but was not found when the tests were administered after all of the lists had been presented.

Thus, as Gallo (2006) has indicated, other processes that are likely more based on the stability of the meaning of the event or thematic representation might need to be conducted in order to explain the persistence of false memories.

Concerning the explanations for the DRM phenomenon, recent and ample evidence supports the intervention of both associative activation and thematic extraction processes. On the one hand, several studies have shown that false memories occur by automatic associative activation that spreads from the studied items to the CIs, without consciously identifying the themes of the lists. For example, (a) the study by Seamon, Lee, et al. (2002), in which the participants rehearsed the words aloud at study, showed that they produced high levels of false memories even for the CIs that were never overtly generated; (b) studies using extreme fast presentation of list words, where the conscious generation of the list theme is unlikely during the study, have shown considerable levels of false memories (McDermott & Watson, 2001; Seamon, Luo, & Gallo, 1998); (c) Cleary and Greene (2004) showed that their participants were more likely to falsely recognise the CIs than unrelated words even when the participants were unable to perceive the briefly flashed words in the lists; and (d) conditions of divided attention at study, lowering the likelihood of extracting the theme of the list, have not been able to eliminate false memories (Dewhurst, Barry, & Holmes, 2005; Dewhurst, Barry, Swannell, Holmes, & Bathurst, 2007; Perez-Mata, Read, & Diges, 2002; Peters et al., 2008). On the other hand, evidence suggests that enhancing the likelihood of extracting the theme word increases false memories. For example, (a) blocking the study words by using semantic themes produces higher levels of false memories than presenting them randomly (Toglia et al., 1999); (b) instructions of relational processing facilitates the identification of the themes of the lists and enhances false memories (McCabe, Presmanes, Robertson, & Smith, 2004); (c) a study by Gallo and Seamon (2004) showed that the likelihood of falsely recognising the CIs increased gradually as more word lists were recalled and perceived in an immediate perception task; and (d) studies using the think-out-loud procedure, in which the participants are asked to say everything they think about while studying a list of items (Lampinen, Meier, Arnal, & Leding, 2005; Lampinen, Ryals, & Smith, 2008; Seamon, Lee, et al., 2002; Seamon, Luo, Schwartz, et al., 2002), have shown that as more CIs are thought about during encoding, more false memories tend to be produced. In sum, the present evidence suggests that false memories can occur by associative activation and thematic extraction. False memories seem to arise by the automatic associative activation that spreads from the studied items to the CI without a conscious processing of the theme word, but it can also be enhanced by the conscious extraction of the thematic representation of the event.

While recent studies (Coane, McBride, Termonen, & Cutting, 2015) have started to explore ways in which the two types of processing, associative and thematic, combine to provide additive effects, it is also crucial to pay attention to studies that have teased apart associative activation and thematic extraction. For example, a study by Carneiro, Garcia-Marques, Fernandez, and Albuquerque (2014) aimed to dissociate these two processes by using lists, with two different CIs in each, one associative critical item (AI) and one thematic critical item (TI), and showed that for the same event, both processes contributed to the false memory explanation. The associative activation played a large role in stimulating the error, showing that AIs, even without the quality of "being good themes", produced high levels of false recognition. On the other hand, TIs even without receiving high automatic spreading activation from the studied items, also showed considerable levels of false recognition. This finding supports the intervention of both automatic activation and thematic extraction in explaining the DRM phenomenon. However, this study and other previous studies also showed that with respect to memory editing, only thematic extraction has a role in the rejection of false memories. Only thematic false recognition was significantly higher under a speeded condition at test than under the self-paced presentation of items of the recognition test (Carneiro, Fernandez, & Dias, 2009; Carneiro et al., 2012; Garcia-Marques, Fernandez, & Albuquerque, 2014).

Using the same double-CI list methodology, the present study aimed to analyse how associative CIs and thematic CIs behave with the passage of time. If the gist trace is highly durable, as fuzzy trace theory proclaims, and if automatic activation decays with time, as was found in lexical decision studies, then thematic false memories should be more stable across time than associative false memories. Two experiments with immediate and delayed recognition test conditions were conducted to address this issue, with one of them manipulating the study context of the to-be-recognised items and the other manipulating the timing of the recognition responses.

Experiment 1

The aim of this experiment was to determine the relative strength of associative and thematic CIs across time by using a recognition test in which the probes consisted of word pairs and in which the response was to be selected from among four alternatives (item 1, item 2, both, or none). This testing procedure was chosen because it involved trials that required the participants to directly contrast the associative and the thematic CIs in each list, also allowing the participants to respond that *none* or *both* items were presented. Considering the two CIs simultaneously rather than sequentially (as in standard recognition tests) makes the overlap in the qualitative characteristics of the two items more salient and, hence, may force the participants to rely on familiarity to a lesser extent and to use a stricter criterion in providing an answer (Dodson & Johnson, 1993).

One group of participants performed this test immediately after the presentation of lists (immediate condition), and another group performed the task one week later (delay condition). Moreover, for 50% of the participants in each group, the CIs were presented within the study lists, providing measures of veridical recognition, and for the remaining 50% of the participants, the CIs were not presented in the study phase, allowing measures of false recognition of the same words across the experiment. The procedure of using omitted/presented CIs in lists has already been employed in forced-choice studies (e.g., Weinstein, McDermott, & Chan, 2010; Westerberg & Marsolek, 2003). Westerberg and Marsolek (2003) showed that in forced-choice tests, participants could better distinguish the presented and omitted words in the lists than the presented and omitted CIs, whereas Weinstein et al. (2010) showed that participants could properly distinguish the studied words from unpresented CIs in an immediate test but not in a delayed test, after a 7-day retention interval. In the present study, the omitted/presented CIs procedure was used not only to analyse the differences in false recognition between AIs and TIs but also to examine whether AIs and TIs are equally well remembered when they are presented within lists, both in an immediate and in a delay condition.

In this experiment, we predicted more stability of false memories across time for TIs than for AIs. If forgetting thematic false memories occurs more slowly than forgetting associative false memories, more false recognition for AI than for TI was expected in an immediate test (i.e., the same pattern of previous experiments, e.g., Carneiro et al., 2014), but no differences or even the opposite pattern was expected in a delayed test.

Method

Participants

A total of 128 Portuguese university students participated in this experiment (M_{age}=21 years, SD = 6.79; 104 females). Sixty-five participants were assigned to the immediate condition (n = 33 in the present CIs condition and n = 32 in the absent CIs condition), and 63 students were assigned to the delay condition (n = 33 in the present CIs condition and n = 30 in the absent CIs condition). All of them received course credit for participating in the experiment.

Design

The experiment followed a 2 (type of CI: AI vs. TI) × 2 (retention interval: immediate vs. delay) × 2 (status of CI: in lists vs. out lists) factorial design, using repeated measures over the first factor.

Material and procedure

Twelve DRM lists already normed for a Portuguese population (Carneiro, Ramos, Costa, Garcia-Marques, & Albuquerque, 2011) were used as the specific target lists in this study. These lists were especially chosen for this study. In a previous identifiability task (Carneiro et al., 2011), we observed that some lists had CIs that easily elicit the themes of the lists, whereas others do not. This was found by presenting the DRM lists to the participants and asking them to generate a single word that best defined the overall theme of each list. For most lists, the theme that was most frequently generated by the participants coincided with the associatively derived CI of the list. However, for a few lists, the participants generated other words as themes with a higher frequency than the associative CIs. As occurred in Carneiro et al. (2014), we selected these special lists by trying to separate associative and thematic processes. Therefore, for each list, two CIs were considered: the associative (AI), derived from free association norms, and the thematic (TI), derived from identifiability norms. For example, for the list comprising the words *dilatory, snail, delayed, unhurried, fast, slug, wimpy, lazy, laziness,* and *calm,* the AI is *slow* and the TI is *speed. Slow* received the highest associative activation from the words of the list (as evidenced by the association norms), and *speed* was the word most defined as the theme of the list (as evidenced by the identifiability norms). The associative and thematic lures of the target lists differed significantly in theme identifiability values, with AIs showing lower percentages of theme identifiability than TIs (M_{AI}= 7.4 vs. M_{TI} = 18.8, $t(11) = 7.08$, $p = .001$) and in MBAS (mean backwards associative strength), with AIs showing higher MBAS than TIs (M_{AI} = .13 vs. M_{TI} = .02, $t(11) = 4.07$, $p = .002$). The two types of items were controlled for similarity in linguistic dimensions, such as length (M_{AI} = 5.6 vs. M_{TI} = 6.4, $t(11) = 1.33$, $p = .21$), frequency in the language (M_{AI} = 1893.8 vs. M_{TI} = 2655.6, $t(11) = .90$, $p = .39$), concreteness (M_{AI} = 5.1 vs. M_{TI} = 4.3, $t(11) = 1.36$, $p = .20$), familiarity (M_{AI} = 5.3 vs. M_{TI} = 5.2, $t(11) = .34$, $p = .74$), and imagery (M_{AI} = 4.6 vs. M_{TI} = 4.4, $t(11) = .48$, $p = .64$).

In both retention interval conditions, immediate and delay, 15 associative lists of 10 words each were presented, with 12 target lists and 3 filler lists (all of them presented in Appendix 1). The 3 filler lists were DRM lists that were not related to the themes of the target lists, and they were presented to serve as material for an initial recognition test in the delay condition.

For 50% of the participants in each condition, the CIs were presented within the study lists. In this case, the fourth and sixth associates of each list were removed and substituted for the AI and the TI in that list (with the order of appearance of the CIs counterbalanced across the participants). For the other 50% of the participants, the CIs were not presented in the study, and the ten strongest associates were presented, as in the standard DRM paradigm.

The 15 lists were auditory and randomly presented at a rate of 2s/word, and the words within each list were kept in the same order for each participant. After each list's presentation, an "End of list" message appeared on the computer screen. The participants had to press the space key to hear the following list.

In the immediate condition, the participants did a four-alternative recognition test in which they were tested in relation to the 12 target lists, immediately after the presentation of the last of the 15 lists. In this condition, memory for the items in the 3 filler lists was not tested. In the delay condition, the participants performed a yes/no recognition test, including 5 old items from each of the 3 filler lists and 15 new words extracted from similar DRM lists not used in this study, immediately after the presentation of last study list, and one week after they performed a surprise four-alternative recognition test for the 12 target lists. This initial test was conducted only to eliminate participants' expectations about a delayed test and to prevent extra practice with the four-alternative recognition test; consequently, these results were not analysed. In the delay condition, to secure their return in a week, the participants were initially informed that to receive the course credit for participating, they had to return one week later to perform other types of tasks.

In the four-alternative recognition test of both retention interval conditions, a pair of items was shown on the computer screen, and the participants had to decide among four given response alternatives: *both items were previously presented, only the left item was previously presented, only the right item was previously presented* or *none of the items was previously presented*.

In either condition, immediate or delay, the recognition test for the target lists consisted of 108 pairs of words (9 pairs for each of the 12 target lists). The items that were tested, for each list, were the AI, the TI, the first associate (S1), the fifth associate (S5) and one unrelated/non-presented item (U) of similar frequency, resulting in nine possible pairs for each list: AI–TI; AI–S1; AI–S5; TI–S1; TI–S5; AI–U; TI–U; S1–U; and S5–U. The order of the presentation of the lists, the order of the pairs per list, and the position of each item in the pair (left or right) were randomly assigned. Note that because both the AI and TI were equally repeated, potential effects of repeated testing would be similar for the two types of critical words.

Results and discussion

Each group of pairs of items was separately analysed.

Pair AI–TI

As the AI–TI pair is the most important for the aim of the present study, it was analysed in detail. The distribution of average *yes* responses (recognition proportions) to the four alternatives (*both, AI, TI,* and *none*) is presented in Figure 1(a). To compare the recognition rates of the AI and TI as a function of retention interval and the presence/absence of the CIs in the study lists (Figure 1(b)), the recognition scores of each item reflected the proportion of times that the studied item was identified, either by an item-specific response or by a *both* response.

With these values, a 2 (type of CI: AI vs. TI) × 2 (retention interval: immediate vs. delay) × 2 (status of CI: in lists vs. out lists) ANOVA, with repeated measures over the first factor, was performed. The results showed three main effects: a main effect of type of CI, $F(1, 124) = 35.39$, MSE = .019, $p < .001$, $\eta_p^2 = .22$, showing that, in general, the AI produced higher levels of recognition than the TI (M_{AI} = .62, SD = .21 vs. M_{TI} = .51, SD = .25); a main effect of retention interval, $F(1, 124) = 37.39$, MSE = .049, $< .001$, $\eta_p^2 = .232$, with

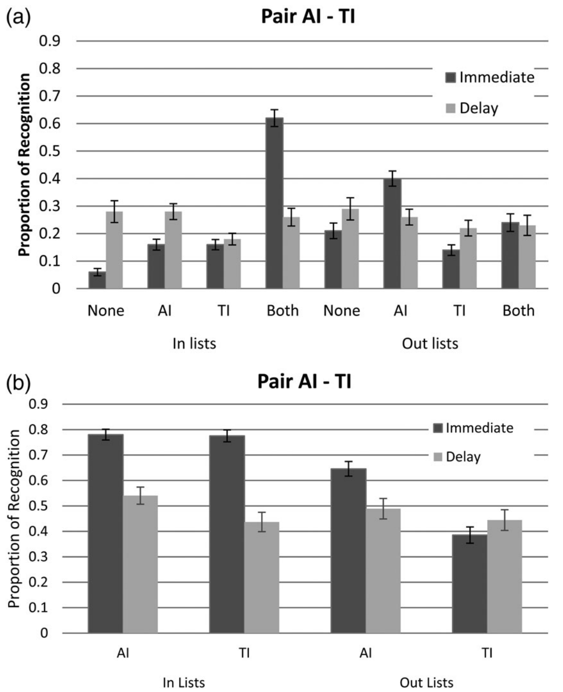

Figure 1. (a) and (b) Proportions of recognition and standard errors for the pair AI–TI (with the four alternatives of response – 1a, and only for the two critical items – 1b), as a function of retention condition (immediate vs. delay) and status of CIs (in lists vs. out lists) (Experiment 1).

higher levels of recognition in the immediate condition than in the delay condition ($M_I = .65$, SD = .21 vs. $M_D = .48$, SD = .21); and another main effect of status of CI, $F(1, 124) = 26.41$, MSE = .049, $p < .001$, $\eta_p^2 = .176$, with more recognition when the CIs were in lists (veridical recognition) than when CIs were omitted (false recognition) ($M_{in} = .63$, SD = .22 vs. $M_{out} = .49$, SD = .21).

The results also showed a significant type of CI × status of CI interaction, $F(1, 124) = 7.98$, MSE = .019, $p = .01$, $\eta_p^2 = .06$, indicating that the difference between AI and TI is higher for false recognition (when CIs were out lists) than for veridical recognition (when CIs were in lists) (out lists: $M_{AI} = 57$, SD = .21 vs. $M_{TI} = .42$, SD = .20, $p < .001$; in lists: $M_{AI} = .66$, SD = .20 vs. $M_{TI} = .61$, SD = .25, $p = .03$), although both differences were significant. A significant retention interval × status of CI interaction also emerged, $F(1, 124) = 18.87$, MSE = .049, $p < .001$, $\eta_p^2 = .132$, revealing that when CIs appeared in the lists, they were significantly more correctly recognised in the immediate condition than one week later ($M_I = .78$, SD = .13 vs. $M_D = .49$, SD = .21, $p < .001$), whereas when CIs were absent, false recognition did not differ across time ($p = .22$).

These interactions were qualified by a significant type of CI × retention interval × status of CI interaction, $F(1, 124) = 20.47$, MSE = .019, $p < .001$, $\eta_p^2 = .142$, revealing that when CIs were omitted from the lists, false recognition for the AI was significantly different across retention conditions ($p = .001$), where more false recognition occurred in the immediate condition than in the delay condition ($M_I = .65$, SD = .16 vs. $M_D = .49$, SD = .22), whereas false recognition for the TI did not differ across retention conditions ($M_I = .39$, SD = .18 vs. $M_D = .44$, SD = .22, $p = .23$). This result seems to suggest that thematic false memories are more stable across time, whereas associative false memories are more vulnerable to time decay.

When CIs were present in the lists, recognition for the both the associative and thematic critical items differed significantly across retention conditions, producing more correct recognitions in the immediate condition than in the delay condition ($M_I = .78$, SD = .12 vs. $M_D = .54$, SD = .19, $p = .00$, for the associative critical item; $M_I = .78$, SD = .13 vs. $M_D = .44$, SD = .22, $p = .00$, for the thematic critical item). This result seems to corroborate previous findings that, in general, veridical memories are more vulnerable to time decay than false memories (McDermott, 1996; Thapar & McDermott, 2001).

This significant three-way interaction between the type of CI, the retention interval and the status of CI also reveals that in the immediate condition, there was a significant difference between the AI and TI when they were omitted from the lists, with the AI being more falsely recognised than the TI ($M_{AI} = .65$, SD = .16 vs. $M_{TI} = .39$, SD = .18, $p < .001$), whereas when the CIs were in the lists, this difference was not significant, with the AI being as correctly recognised as the TI ($M_{AI} = .78$, SD = .12 vs. $M_{TI} = .78$, SD = .13). On the one hand, this result corroborates the previous finding (Carneiro et al., 2014) that the AI is more prone to be falsely recognised than the TI in an immediate condition. On the other hand, it suggests that both items are equally prone to elicit true memories in tests that immediately follow the presentation of lists. In the delay condition, the difference in false recognition between the AI and TI was not significant ($M_{AI} = .49$, SD = .22 vs. $M_{TI} = .44$, SD = .22, $p = .22$), whereas veridical recognition significantly differed between the CIs, with the AI producing higher levels of correct recognition than the TI ($M_{AI} = .54$, SD = .19 vs. $M_{TI} = .44$, SD = .22, $p = .003$).

In short, associative false memories were higher than thematic false memories in the immediate condition; however, they decayed more rapidly over time, reaching the same levels as thematic false memories within one week. When the two types of CIs were presented in the lists, they produced the same amount of correct recognition in the immediate condition; however, after a one-week delay, the AI produced higher levels of correct recognition than the TI, a result that may be related to the possibility that the presence of the AI's studied associates in the test leads to extra activation of the AI (Coane & McBride, 2006; Meade, Watson, Balota, & Roediger, 2007).

Pairs AI–S and TI–S

Although the AI–TI pair is the most imperative for the aim of the present study, it is also important to analyse the behaviour of these CIs when they are contrasted with other items, such as the studied items (see Figure 2). ANOVAs with similar factorial designs were performed for the following pairs: AI–S1, TI–S1, AI–S5, and TI–S5. For the

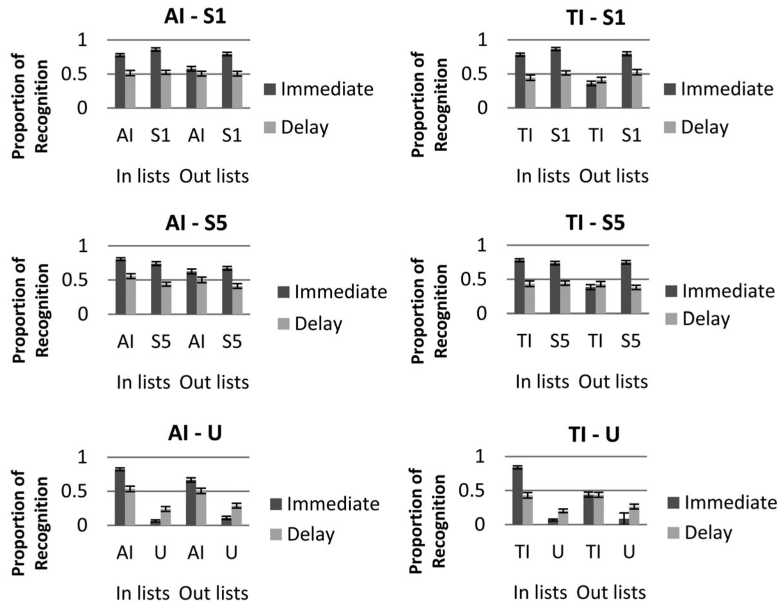

Figure 2. Proportions of recognition and standard errors for the pairs AI–S1, TI–S1, AI–S5, TI–S5, AI–U, and TI–U, as a function of retention condition (immediate vs. delay) and status of CIs (in lists vs. out lists) (Experiment 1).

sake of simplicity, the detailed statistical analyses for these pairs are presented in Appendix 2. For all of these pairs, the results of these analyses showed that recognition in the immediate condition was generally higher than in the delay condition and that veridical recognition was generally higher than false recognition. In most cases, the studied items produced higher levels of recognition than the AI and TI, except for the pair IA–S5, for which the reverse was found.

Regarding the comparison between the false (AI and TI) and veridical recognition of the first associate (S1), the results corroborate the findings of Weinstein et al. (2010). False memories (for both the AI and TI) were forgotten more slowly than veridical memories, meaning that accuracy was above chance in the immediate test (S > AI; S > TI) but did not differ from chance performance in the delay test (S = AI; S = TI).

Pairs AI–U and TI–U

The analyses of both CIs with unrelated items (distractors) revealed a similar pattern for both the AI and TI: the veridical or false recognition of the AI or IT was always higher than the incorrect recognition of unrelated items, in both the immediate and delay conditions (Figure 2) (the statistical results are presented in Appendix 3). In both pairs, the incorrect recognition of distractors increased in the delay condition, meaning that memory was less accurate after one week. However, regarding the CIs, the same pattern as already found for the other pairs was observed: false recognition of the AI decreased with time, whereas that of the TI remained stable.

Summing up all of the results, we can likely conclude that with the one-week interval, thematic false memories were the only memories that showed stability. Associative false memories, as veridical memories, decreased, and the memory of unrelated distractors increased across that time span. However, after one week, veridical memories and associative and thematic false memories reached similar levels of recognition.

Experiment 2

Experiment 1 showed that associative false memories decrease across time, whereas thematic false memories are similar after a one-week delay. Note, however, that associative false memories were significantly higher than thematic false memories in general, although in the delay condition the difference between associative and thematic false memories was not significant. This result could be interpreted to indicate that associative false memories are more vulnerable to time decay, whereas thematic false memories remain stable across that time span. However, this account could not be the entire story. Considering the processes that are responsible for the occurrence of a false memory, the pattern of results not only might be due to differences in associative activation (MBAS), with words' list eliciting less activation to critical TIs, but also could reflect the different susceptibility of each item to monitoring operations. In fact, previous studies (Carneiro et al., 2014) have shown that AIs and TIs differ in their behaviour during memory editing, with thematic items seeming to be more affected by monitoring processes than associative items. Moreover, it is known that monitoring is not as effective in a delayed test as in an immediate test (Seamon et al., 2006). Thus, it could be the case that thematic false memories were lower than associative false memories in the immediate test because they received less activation from the words' list and because they were more easily edited out in the immediate test. According to this interpretation, the different courses of the AI and TI across time could also imply that they are differentially susceptibility to monitoring.

To explore this possibility, and to better understand the pattern of results of Experiment 1, we conducted the present experiment to study false memories under conditions permitting the operation of an editing mechanism (standard retrieval condition) and under conditions preventing it (speeded retrieval condition). If mainly the monitoring explanation is underlying the pattern of results and if in fact the monitoring of thematic items is operating in the immediate condition and not in the delay condition, we expect to find a decrease in thematic false recognition across a one-week delay when the participants are tested in a condition of speeded retrieval but not when monitoring is permitted (i.e., in the standard retrieval condition). Comparing the results using both types of retrieval, we can more accurately conclude whether the interpretation that only associative false memories decay rapidly across time is a better explanation for the present results or whether monitoring processes also play a role in the explanation of this finding.

Method

Participants

A total of 117 Portuguese university students participated in this experiment (M_{age}=23 years, SD = 5.36; 68 female). Fifty-nine participants were assigned to the immediate condition ($n = 28$ in the speeded retrieval condition and $n = 31$ in the self-paced retrieval condition), and 58 participants were assigned to the delay condition ($n = 28$ in the speeded retrieval condition and $n = 30$ in the self-paced retrieval condition). All of the participants were native Portuguese speakers and received vouchers to participate in the study.

Design

The experiment followed a 2 (type of CI: AI vs. TI) × 2 (retention interval: immediate vs. delay) × 2 (retrieval time: self-paced vs. speeded) factorial design with repeated measures over the first factor.

Material and procedure

Ten target lists and the same three filler lists used in Experiment 1 were used in Experiment 2. The procedure of the presentation of the lists was the same as in the previous experiment, with the important exception that only 10 associates in each list were presented (the critical items were always absent). As in Experiment 1, in the immediate condition, memory for the 10 target lists was tested immediately after the presentation of the lists; moreover, in the delay condition, memory for the filler lists was tested on the first day, and memory for the 10 target lists was tested one week later.

Because of the adopted procedure of speeded response, in this experiment, the participants were submitted to a standard recognition test instead of a four-alternative test. Each item was shown individually on the computer screen, and the participants had to decide whether the item had been previously presented (*old*) or not (*new*). In the self-paced condition, the participants did not have a time limit for giving their responses, whereas in the speeded condition, the participants were instructed to give their responses quickly, not exceeding the time limit of 800 ms (400 ms for each word presentation and 400 ms for a response). Whenever the response exceeded this time limit, a written message of "try to be faster" appeared on the computer screen.

The recognition test for the filler lists, which was administered only on the first day of the delay condition and in a self-paced manner, consisted of 30 randomly presented words (5 old items from each of the 3 filler lists and 15 new words). In both conditions, immediate and delay, the recognition test for the target lists consisted of 68 words, which were randomly presented: 20 words from the selection of 2 studied words per target list (positions 1 and 8); 10 critical associative items of the presented target lists; 10 critical thematic items of the presented target lists; and 28 unrelated words (drawn from the list items, associative critical items and thematic critical items of 7 other DRM lists). The best way to construct the recognition test would be to counterbalance the items that served as targets or CIs and the items that served as distractors (target-controls and CIs-controls) across subjects. In the present study, such a procedure was not practical because we would need to find more lists (at least 20) that fit in to the 2 CIs-list criteria: AIs showing lower percentages of thematic identifiability than TIs and AIs showing higher MBAS than TIs. However, the strategy of using targets and critical lures (AIs and TIs) from other unrelated lists (distractors) is a good approximation to the ideal and has already been employed in previous studies (e.g., Carneiro et al., 2014).

Results and discussion

Figure 3 displays the proportion of recognition for the targets, and Figure 4 displays the proportion of recognition for the associative and thematic CIs in the two retention conditions (immediate vs. delay) and for the two types of retrieval time (speeded vs. standard). The analyses were separately performed for the targets and critical items.

A 2 (retention interval: immediate vs. delay) × 2 (retrieval time: self-paced vs. speeded) ANOVA was performed for targets (veridical recognition). The results showed two main effects: one of retention interval, $F(1, 113) = 90.91$, MSE = .029, $p < .001$, $\eta_p^2 = .45$, and another one of retrieval time, $F(1, 113) = 7.94$, MSE = .029, $p = .006$, $\eta_p^2 = .07$, showing that the participants recognised more correct items in the immediate condition than in the delay condition ($M_I = .72$, SD = .17 vs. M_D=.42, SD = .18) and in the

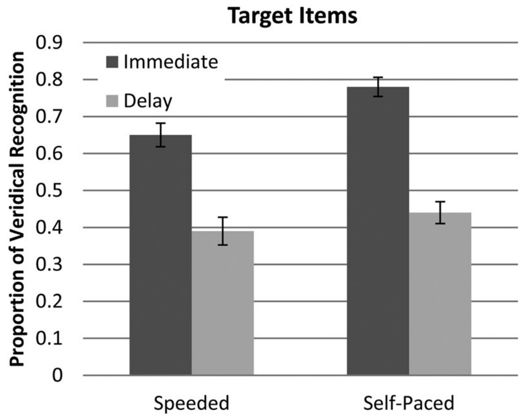

Figure 3. Proportions of recognition and standard errors for targets, as a function of retention condition (immediate vs. delay) and retrieval time (speeded vs. self-paced) (Experiment 2).

self-paced condition than in the speeded condition (M_{SP} = .61, SD = .23 vs. M_S = .52, SD = .23).

To analyse the recognition of critical items (false recognition), a 2 (type of CI: AI vs. TI) × 2 (retention interval: immediate vs. delay) × 2 (retrieval time: self-paced vs. speeded) ANOVA, with repeated measures over the first factor, was performed. The following results were obtained: a main effect of type of CI, $F(1, 113) = 56.13$, MSE = .021, $p < .001$, $\eta_p^2 = .33$, with AIs producing higher levels of false recognition than TIs (M_{AI}=.52, SD = .22 vs. M_{TI} = .37, SD = .22); a main effect of retention interval, $F(1, 113) = 20.80$, MSE = .06, $p < .001$, $\eta_p^2 = .16$, with more false recognition in the immediate condition than in the delay condition (M_I = .52, SD = .20 vs. M_D = .37, SD = .21); and a main effect of retrieval time, $F(1, 113) = 4.54$, MSE = .06, $p = .04$, $\eta_p^2 = .04$, with the speeded condition generally producing higher levels of false recognition than the self-paced condition (M_S = .48, SD = .24 vs. M_{SP} = .41, SD = .20).

The results also showed a significant critical item × retrieval time interaction, $F(1,113) = 6.39$, MSE = .021, $p = .013$, $\eta_p^2 = .05$, revealing that for the AI, the proportion

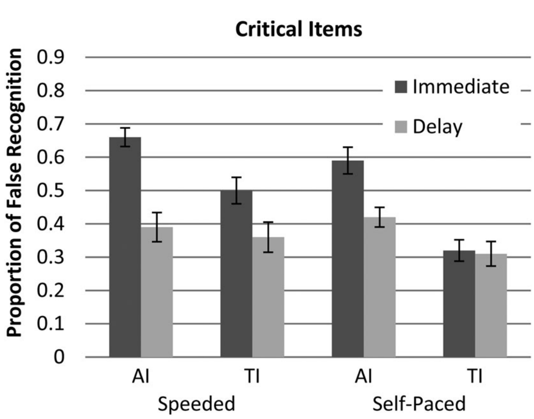

Figure 4. Proportions of recognition and standard errors for critical items (AI and TI), as a function of retention condition (immediate vs. delay) and retrieval time (speeded vs. self-paced) (Experiment 2).

of false recognition was not significantly different in the two retrieval time conditions, whereas for the TI, false recognition was significantly higher in the speeded condition than in the self-paced condition (M_S = .43, SD = .23 vs. M_{SP} = .31, SD = .19, $p = .003$). This result seems to corroborate the previous finding (Carneiro et al., 2014) that thematic items are more likely to be rejected than associative items in general.

Another significant type of CI × retention interval interaction was found, $F(1, 113) = 15.60$, MSE = .021, $p < .001$, $\eta_p^2 = .12$, showing that the difference between the immediate and the delay condition was significant only for the AI (M_I = .62, SD = .19 vs. M_D = .40, SD = .20, $p < .001$), whereas this difference was not significant for the TI (M_I = .41, SD = .22 vs. M_D = .34, SD = .22, $p = .07$). This significant interaction seems to corroborate the results in Experiment 1 and indicates that AIs are more vulnerable to time decay than TIs. The interaction between retention interval and retrieval time was not significant.

Although the three-way interaction between the type of CI, retention interval and retrieval time did not reach significance ($F < 1$), planned comparisons testing our specific hypothesis indicated that, as expected, there was no significant difference for false recognition of the TI in the self-paced condition between the immediate and delayed test, $t(113) = 0.12$, $p > .05$; however, in the speeded condition, the false recognition of the TI was significantly affected by the retention time, $t(113) = 2.46$, $p = .02$, with higher levels of false recognition in the immediate condition than in the delay condition (M_I = 50, SD = .21 vs. M_D = 36, SD = .24). This result is consistent with our initial hypothesis that the stability of the TI across time might be explained by its vulnerability to monitoring processes.

For completeness, and to compare the time course of the studied and critical items, a 3 (type of item: studied vs. associative vs. thematic) × 2 (retention interval: immediate vs. delay) ANOVA with repeated measures over the first factor was performed. The results showed a main effect of type of item, $F(2, 230) = 53.06$, MSE = .023, $p < .001$, $\eta_p^2 = .32$, revealing that, in general, the participants correctly recognised more studied items (M_{SI} = .57, SD = .23) than falsely recognised critical items (M_{AI} = .51, SD = .22; M_{TI} = .37, SD = .22). As the previous analysis showed, in general, the AI was more falsely recognised than the TI ($p < .001$). Unsurprisingly, the immediate condition produced higher levels of recognition than the delay condition, ($F(1, 115) = 48.89$, MSE = .069, $p < .001$, $\eta_p^2 = .30$; M_I = .58, SD = .19 vs. M_D = .39, SD = .20). More interestingly, a significant type of item × retention interval interaction was found, $F(2, 230) = 17.54$, MSE = .023, $p < .001$, $\eta_p^2 = .13$, showing a significant decrease in retention between the immediate and delay conditions for all of the items ($p < .001$), except for the TI ($p = .09$). This interaction also revealed that whereas in the immediate condition, the studied items were more recognised than the unpresented associative CIs (M_{SI} = .72, SD = .17 vs. M_{AI}

= .62, SD = .19), in the delay condition, there was not a significant difference between these two types of items (M_{SI} = .42, SD = .18 vs. M_{AI} = .40, SD = .20). This observation seems to indicate that with a one-week delay, veridical recognition decays to false recognition levels, as occurred in Experiment 1. Although recognition of AIs also decays, veridical recognition and false recognition reached similar levels when the participants were tested one week after the presentation of the lists.

Mean recognition of unrelated items was low (M_U = .17, SD = .20). As expected, unrelated items were more falsely recognised in the speeded condition than in the self-paced (M_S = .26, SD = .24 vs. M_{SP} = .09, SD = .12) and in the delay condition than in the immediate (M_I = .15, SD = .20 vs. M_D = .19, SD = .20).

In sum, the results of Experiment 2 corroborate the results of Experiment 1: (a) in the immediate condition, associative false memories were higher than thematic false memories; (b) whereas veridical recognition and associative false memories decay after a one-week delay, thematic false memories are stable. The present experiment seems to suggest that source-monitoring judgements play a greater role in immediate testing than in delayed testing. Thus, eventually, the stability of the thematic items might also be due to its higher susceptibility to monitoring in immediate testing, which corroborates the results found by Carneiro et al. (2014).

General discussion

The two experiments, which used different types of recognition tests, showed a consistent pattern of results: false recognition of AIs decreased with a one-week delay, whereas false recognition of TIs was similar in the immediate and delay conditions. This seems to indicate that AIs and TIs are different by nature.

In Experiment 1, in which the AI and TI were directly contrasted, the participants chose more often the AI than the TI in the immediate condition; however, there was no preference in the delay test after one week. The advantage of AIs in an immediate test cannot be attributed to the higher distinctiveness of these items in relation to TIs, because when they were both presented, they produced similar levels of veridical recognition.

Given these results, one can conclude that associative false memories are more vulnerable to decay, whereas thematic false memories remain stable across one week. Remember, however, that because the AIs in our lists were not "good themes" for their lists, false recognition of these items would mostly result from automatic activation, which is influential in an immediate condition but does not persist across time. In the case of TIs, however, false recognition mostly depends on them being "good theme identifiers" for their lists, a quality that is rather resistant to the passage of time. However, if we also consider the editing processes, another explanation could be provided for these results: the stability of false thematic recognition across different retention interval conditions could be due to the fact that in the immediate condition, TIs were more susceptible to monitoring than AIs, a result that has already been found by Carneiro et al. (2014). Because monitoring is usually ineffective in a condition with a one-week delay, the apparent temporal stability of TI recognition could be the result of an effective monitoring of TIs in the immediate condition. To analyse the merit of such an explanation, in Experiment 2, a condition in which monitoring is prevented (by speeded retrieval) was included in the design. The comparison between the results obtained with self-paced retrieval and with speeded retrieval, in immediate and delay conditions, allowed us to analyse the influence of monitoring processes. On the one hand, this experiment showed that in both conditions of retention, recognition of TIs was lower in the self-paced condition than in the speeded condition. This result corroborated the previous finding of Carneiro et al. (2014), showing that TIs were more prone to be edited out than AIs in the self-paced condition (although we expected to find significantly lower thematic false recognition in the immediate than in the delay condition of the self-paced condition because of the inefficiency of monitoring in the delay condition).

On the other hand, this experiment showed that in both conditions of retrieval, false recognition of AIs decreased with a one-week delay, whereas false recognition of TIs showed similar levels after such a time span. However, only when monitoring was allowed, as in the self-paced condition, did the recognition of the TI remain stable. When monitoring is prevented (speeded condition), TI recognition decreased across retention conditions, leading us to conclude that as hypothesised, monitoring also contributes to the explanation for the distinct time courses of the AI and TI.

In the present study, the general conclusion that associative false memories, although initially more frequent, decay more rapidly across time than thematic false memories is consistent with the results of Matzen, Taylor, and Benjamin (2011). In that study, the authors found a more rapid loss of access to conjunction lures, defined as having high orthographic overlap with the studied items, than to semantic lures, defined as having high semantic overlap with the studied items. Because TIs, in comparison with AIs, are likely to have a higher thematic overlap with the studied items, we can speculate that the results between the two experiments have similar explanations, considering the associative/gist accounts.

Concerning the two dominant theories of false memories, these results provide support mainly for the activation-monitoring theory (Roediger et al., 2001). The results that replicated the findings of Carneiro et al. (2014) by showing higher levels of associative false memories than thematic false memories and the intervention of monitoring processes in the immediate condition support associative activation theories. However, one can argument that the result that showed a higher apparent stability of thematic false memories than of associative false memories across time supports the assumption of fuzzy trace theory that gist is a durable trace (Brainerd & Reyna, 2005).

Together, these results suggest that immediately after stimulus presentation, associative activation is the most important factor for false memory formation; however, with time, this activation fades and starts yielding to thematic extraction. Nevertheless, it is important to note that in the condition with a one-week delay, the level of thematic false memories was never, in the two experiments, higher than that of associative false memories. In that respect, additional investigations should determine whether thematic critical items could reach higher levels of false recognition than associative critical items if the retention interval was higher (e.g., one month or longer).

In this study, we hypothesised that "resistance-to-decay" could be a good candidate and, more specifically, that TIs (reflecting gist) would be more likely to persist over time than AIs (reflecting spreading activation). The findings reported in this manuscript, showing a differential effect of the retention interval on the two types of critical words, are consistent with this hypothesis. What obviously requires more investigation is the nature of the hypothesised mechanisms. Thus, it would make sense to search for the emergence of recognition reversal effects with more extensive manipulations of the retention intervals or to further explore the interactions of the retention interval with other potentially relevant factors, such as specific levels of BAS and the degree of the items' thematic representation (i.e., beyond a dichotomic manipulation of these variables). Still, the findings of the experiments reported in this manuscript can be taken as a relevant move in the right direction, providing an important first step in a potentially important new line of research.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by Foundation for Science and Technology, Portugal [PTDC/PSI-PCO/112466/2009].

References

Anderson, J. R. (1983). A spreading activation theory of memory. *Journal of Verbal Learning and Verbal Behavior*, *22*(3), 261–295. doi:10.1016/S0022-5371(83)90201-3

Brainerd, C. J., & Reyna, V. F. (2005). *The science of false memory*. New York, NY: Oxford University Press.

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*(2), 115–129. doi:10.3758/MC.37.2.115

Carneiro, P., Fernandez, A., Diez, E., Garcia-Marques, L., Ramos, T., & Ferreira, M. B. (2012). "Identify-to-reject": How to avoid false memories in the DRM paradigm. *Memory & Cognition*, *40*(2), 252–265. doi:10.3758/s13421-011-0152-6

Carneiro, P., Garcia-Marques, L., Fernandez, A., & Albuquerque, P. (2014). Both associative activation and thematic extraction count, but thematic false memories are more easily rejected. *Memory*, *22*, 1024–1040. doi:10.1080/09658211.2013.864680

Carneiro, P., Ramos, T., Costa, R. S., Garcia-Marques, L., & Albuquerque, P. (2011). Identificabilidade dos temas de listas formadas por associação retrógrada (backward): Contributo para o estudo das memórias falsas. *Laboratório de Psicologia*, *9*(1), 23–34. doi:10.14417/lp.634

Cleary, A. M., & Greene, R. L. (2004). True and false memory in the absence of perceptual identification. *Memory*, *12*(2), 231–236. doi:10.1080/09658210244000577

Coane, J. H., & McBride, D. M. (2006). The role of test structure in creating false memories. *Memory & Cognition*, *34*(5), 1026–1036. doi:10.3758/BF03193249

Coane, J. H., McBride, D. M., Termonen, M., & Cutting, J. C. (2015). Categorical *and* associative relations increase false memory relative to purely associative relations. *Memory & Cognition*, *44*(1), 37–49. doi:10.3758/s13421-015-0543-1

Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review*, *82*(6), 407–428. doi:10.1037/0033-295X.82.6.407

Cotel, S. C., Gallo, D. A., & Seamon, J. G. (2008). Evidence that nonconscious processes are sufficient to produce false memories. *Consciousness & Cognition*, *17*(1), 210–218. doi:10.1016/j.concog.2007.01.009

Deese, J. (1959). On the prediction of occurrence of particular verbal intrusions in immediate recall. *Journal of Experimental Psychology*, *58*(1), 17–22. doi:10.1037/h0046671

Dewhurst, S. A., Barry, C., & Holmes, S. J. (2005). Exploring the false recognition of category exemplars: Effects of divided attention and explicit generation. *European Journal of Cognitive Psychology*, *17*(6), 803–819. doi:10.1080/09541440540000013

Dewhurst, S. A., Barry, C., Swannell, E. R., Holmes, S. J., & Bathurst, G. L. (2007). The effects of divided attention on false memory depend on how memory is tested. *Memory & Cognition*, *35*(4), 660–667. doi:10.3758/BF03193304

Dodson, C. S., & Johnson, M. K. (1993). Rate of false source attributions depends on how questions are asked. *American Journal of Psychology*, *106*(4), 541–557. doi:10.2307/1422968

Gallo, D. A. (2006). *Associative illusions of memory: False memory research in DRM and related tasks*. New York, NY: Psychology Press.

Gallo, D. A. (2010). False memories and fantastic beliefs: 15 years of the DRM illusion. *Memory & Cognition, 38*, 833–848. doi:10.3758/MC.38.7.833

Gallo, D. A., & Seamon, J.G. (2004). Are nonconscious processes sufficient to produce false memories? *Consciousness and Cognition*, *13*(1), 158–168. doi:10.1016/j.concog.2003.09.001

Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. *Psychological Bulletin*, *114*(1), 3–28. doi:/10.1037/0033-2909.114.1.3

Lampinen, J. M., Meier, C., Arnal, J. A., & Leding, J. K. (2005). Compelling untruths: Content borrowing and vivid false memories. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *31*(5), 954–963. doi:10.1037/0278-7393.31.5.954

Lampinen, J. M., Ryals, D. B., & Smith, K. (2008). Compelling untruths: The effect of retention interval on content borrowing and vivid false memories. *Memory*, *16*(2), 149–156. doi:10.1080/09658210701839277

Matzen, L. E., Taylor, E. T., & Benjamin, A. S. (2011). Contributions of familiarity and recollection rejection to recognition: Evidence from the time course of false recognition for semantic and conjunction lures. *Memory*, *19*(1), 1–16. doi:10.1080/09658211.2010.530271

McCabe, D. P., Presmanes, A. G., Robertson, C. L., & Smith, A. D. (2004). Item-specific processing reduces false recognition. *Psychonomic Bulletin & Review*, *11*(6), 1074–1079. doi:10.3758/BF03196739

McDermott, K. B. (1996). The persistence of false memories in list recall. *Journal of Memory and Language*, *35*(2), 212–230. doi:10.1006/jmla.1996.0012

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 & Language*, *45*(1), 160–176. doi:10.1006/jmla.2000.2771

Meade, M. L., Watson, J. M., Balota, D. A., & Roediger, H. L. (2007). The roles of spreading activation and retrieval mode in producing false recognition in the DRM paradigm. *Journal of Memory and Language*, *56*(3), 305–320. doi:10.1016/j.jml.2006.07.007

Neely, J. H. (1977). Semantic priming and retrieval from lexical memory: Roles of inhibition less spreading activation and limited-capacity attention. *Journal of Experimental Psychology: General*, *106*(3), 226–254. doi:10.1037/0096-3445.106.3.226

Perez-Mata, M. N., Read, J. D., & Diges, M. (2002). Effects of divided attention and word concreteness on correct recall and false memory reports. *Memory*, *10*(3), 161–177. doi:10.1080/09658210143000308

Peters, M. J. V., Jelicic, M., Gorski, B., Sijstermans, K., Giesbrecht, T., & Merckelbach, H. (2008). The corrective effects of warning on false memories in the DRM paradigm are limited to full attention conditions. *Acta Psychologica*, *129*(2), 308–314. doi:10.1016/j.actpsy.2008.08.007

Reyna, V. F., & Brainerd, C. J. (1995). Fuzzy-trace theory: An interitem synthesis. *Learning and Individual Differences*, *7*(1), 1–75. doi:10.1016/1041-6080(95)90031-4

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, DC: 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, & Cognition*, *21*(4), 803–814. doi:10.1037/0278-7393.21.4.803

Seamon, G. J., Lee, A. I., 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 Deese, Roediger, and McDermott procedure. *Psychological Science*, *13*(6), 526–531. doi:10.1111/1467-9280.00492

Seamon, J. G., Berko, J. R., Sahlin, B., Yu, Y. L., Colker, J. M., & Gottfried, D. H. (2006). Can false memories spontaneously recover? *Memory*, *14*(4), 415–423. doi:10.1080/09658210500420725

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*(1), 20–26. doi:10.1111/1467-9280.00004

Seamon, J. G., Luo, C. R., Kopecky, J. J., Price, C. A., Rothschild, L., Fung, N. S., & Schwartz, M. A. (2002). Are false memories more difficult to forget than accurate memories? The effect of retention interval on recall and recognition. *Memory & Cognition*, *30*(7), 1054–1064. doi:10.3758/BF03194323

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.

Seamon, J. G., Luo, C. R., Schwartz, M. A., Jones, K. J., Lee, D. M., & Jones, S. J. (2002). Repetition can have similar or different effects on accurate and false recognition. *Journal of Memory and Language*, *46*(2), 323–340. doi:10.1006/jmla.2001.2811

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*(3), 424–432. doi:10.3758/BF03196393

Toglia, M. P., Neuschatz, J. S., & Goodwin, K. A. (1999). Recall accuracy and illusory memories: When more is less. *Memory*, *7*(2), 233–256. doi:10.1080/ 741944069

Tse, C.-S., & Neely, J. H. (2005). Assessing activation without source monitoring in the DRM paradigm. *Journal of Memory and Language*, *53*(4), 532–550. doi:10.1016/j.jml.2005.07.001

Weinstein, Y., McDermott, K. B., & Chan, J. C. K. (2010). True and false memories in the DRM paradigm on a forced choice test. *Memory*, *18*(4), 375–384. doi:10.1080/09658211003685533

Westerberg, C. E., & Marsolek, C. J. (2003). Sensitivity reductions in false recognition: A measure of false memories with stronger theoretical implications. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*(5), 747–759. doi:10.1037/0278-7393.29.5.747

Appendix 1. Target lists used in Experiment 1 (translated for English language). Experiment 2 used only nine of these lists (three first lists excluded)

IA	Face	Telephone	Shield	God	Hot	Anger
IT	Woman	Friendship	Fatherland	Belief	Winter	Emotions
P1	Visage	Communication	Coin	Faith	Heater	Rage
P2	Eyes	Speak	Euro	Religion	Blankets	Fury
P3	Pretty	Talk	Money	Gift	Fireplace	Cholera
P4	Countenance	Cell phone	Protection	Christ	Temperature	Hate
P5	Tails	Friends	Defence	Pray	Summer	Nerves
P6	Person	Touch	Portugal	Jesus	Scarf	Aggressiveness
P7	Beauty	Numbers	Sword	Atheist	Sweater	Mad
P8	Better half	Contact	Banknote	Church	Coat	Annoying
P9	Expression	Connection	War	Pagan	Fire	Evil
P10	Smile	Call	Weapon	Believer	Cold	Despair
IA	Write	White	Dirty	Slow	Chair	Tall
IT	School	Dentist	Garbage	Speed	Room	Height
P1	Pen	Tooth	Pig	Sluggish	Seat	Short
P2	Read	Peace	Smudge	Snail	Table	Building
P3	Paper	Red	Polluted	Lingering	Backseat	Stepladder
P4	Pencil	Paper	Pollution	Slowly	Couch	Peak
P5	Letter	Cotton	Filth	Fast	Stool	Skinny
P6	Ballpoint	Pure	Dump	Slug	Furniture	Giant
P7	Notebook	Black	Disgust	Coddle	Wood	Hill
P8	Letters	Doctor	Residuals	Lazy	Object	Sound
P9	Sheet	Colour	Sanitary	Laziness	Class	Unachievable
P10	Hand	Bone	Dirtiness	Calm	Chain	Big

Portuguese lists.

IA	Cara	Telefone	Escudo	Deus	Quente	Raiva
IT	Mulher	Amizade	Pátria	Crença	Inverno	Emoções
P_1	Face	Comunicação	Moeda	Fé	Aquecedor	Ira
P_2	Olhos	Falar	Euro	Religião	Cobertores	Fúria
P_3	Bonita	Conversa	Dinheiro	Dádiva	Lareira	Cólera
P_4	Rosto	Telemóvel	Protecção	Cristo	Temperatura	Ódio
P_5	Coroa	Amigos	Defesa	Rezar	Verão	Nervos
P_6	Pessoa	Toque	Portugal	Jesus	Cachecol	Agressividade
P_7	Beleza	Números	Espada	Ateu	Camisola	Zangado
P_8	Metade	Contacto	Notas	Igreja	Casaco	Irritante
P_9	Expressão	Ligação	Guerra	Pagão	Fogo	Maldade
P_{10}	Sorriso	Chamada	Arma	Crente	Frio	Desespero
IA	Escrever	Branco	Sujo	Lento	Cadeira	Alto
IT	Escola	Dentista	Lixo	Velocidade	Sala	Altura
P_1	Caneta	Dente	Porco	Vagaroso	Sentar	Baixo
P_2	Ler	Paz	Borrão	Caracol	Mesa	Prédio
P_3	Papel	Tinto	Poluído	Demorado	Assento	Escadote
P_4	Lápis	Papel	Poluição	Devagar	Sofá	Pico
P_5	Carta	Algodão	Porcaria	Rápido	Banco	Magro
P_6	Esferográfica	Puro	Lixeira	Lesma	Móvel	Gigante
P_7	Caderno	Preta	Nojo	Molengão	Madeira	Monte
P_8	Letras	Médico	Resíduos	Preguiçoso	Objecto	Som
P_9	Folha	Cor	Higiénico	Preguiça	Aulas	Inatingível
P_{10}	Mão	Osso	Sujidade	Calmo	Cadeia	Grande

Appendix 2

Results for the Pair AI–S1: The following results were obtained: a main effect of type of item, $F(1, 124) = 21.73$, MSE = .018, $p < .001$, $\eta_p^2 = .15$, ($M_{S1} = .67$, SD = .23 vs. $M_{AI} = .60$, SD = .21); a main effect of retention time, $F(1, 124) = 87.88$, MSE = .042, $p < .001$, $\eta_p^2 = .42$, ($M_I = .75$, SD = .17 vs. $M_D = .51$, SD = .19); a main effect of status of CI $F(1, 124) = 8.35$, MSE = .042, $p = .005$, $\eta_p^2 = .06$, ($M_{in} = .67$, SD = .23 vs. $M_{out} = .60$, SD = .21); a significant type of item × retention interval interaction, $F(1, 124) = 18.98$, MSE = .018, $p < .001$, $\eta_p^2 = .13$; and a significant retention interval × status of CI interaction $F(1, 124) = 5.37$, MSE = .042, $p = .02$, $\eta_p^2 = .04$. The interactions were qualified by a significant type of item × retention interval × status of CI interaction, $F(1, 124) = 4.63$, MSE = .018, $p = .03$, $\eta_p^2 = .04$, indicating that when CIs were omitted, false recognition for the AI did not significantly differ across retention interval conditions ($p = .12$), whereas the S1 was significantly more recognised in the immediate condition than one week later ($M_I = .79$, SD = .14 vs. $M_D = .51$, SD = .19, $p < .001$). Interestingly, when the CIs were presented, both items were significantly more correctly recognised in the immediate condition than one week later ($M_I = .78$, SD = .12 vs. $M_D = .52$, SD = .22, $p = .00$, for the AI; $M_I = .86$, SD = .14 vs. $M_D = .53$, SD = .18, $p < .001$, for the S1).

Results for the Pair TI–S1: The following results were obtained: a main effect of type of item, $F(1, 124) = 75.25$, MSE = .026, $p < .001$, $\eta_p^2 = .38$, ($M_{S1} = .68$, SD = .23 vs. $M_{TI} = .50$, SD = .26); a main effect of retention interval, $F(1, 124) = 79.16$, MSE = .042, $p < .001$, $\eta_p^2 = .39$, ($M_I = .70$, SD = .21 vs. $M_D = .48$, SD = .21); a main effect of status of CI, $F(1, 124) = 25.50$, MSE = .042, $p < .001$, $\eta_p^2 = .17$, ($M_{in} = .65$, SD = .25 vs. $M_{out} = .53$, SD = .22); a significant type of item × retention interval interaction, $F(1, 124) = 17.11$, MSE = .026, $p < .001$, $\eta_p^2 = .12$; a significant type of item × status of CI interaction, $F(1, 124) = 23.87$, MSE = .026, $p < .001$, $\eta_p^2 = .16$; and a significant retention interval × status of CI interaction, $F(1, 124) = 21.07$, MSE = .042, $p < .001$, $\eta_p^2 = .15$.

These results were qualified by a type of item × retention interval × status of CI interaction, $F(1, 124) = 14.63$, MSE = .026, $p < .001$, $\eta_p^2 = .11$, revealing that when CIs were omitted from the lists, false recognition for the TI did not differ significantly across time ($p = .30$), whereas the S1 was significantly more recognised in the immediate test than one week later ($M_I = .80$, SD = .17 vs. $M_D = .53$, SD = .21, $p < .001$). When CIs were presented, both items were significantly more correctly recognised in the immediate condition than in the delayed condition ($M_I = .78$, SD = .13 vs. $M_D = .44$, SD = .23, $p < .001$, for the TI; $M_I = .87$, SD = .13 vs. $M_D = .52$, SD = .18, $p < .001$, for the S1).

Results for the Pair AI–S5: The following results were obtained: a main effect of type of item, $F(1, 124) = 11.08$, MSE = .018, $p = .001$, $\eta_p^2 = .08$, ($M_{AI} = .62$, SD = .22 vs. $M_{S5} = .57$, SD = .22); a main effect of retention interval, $F(1, 124) = 78.34$, MSE = .045, $p < .001$, $\eta_p^2 = .39$, ($M_I = .71$, SD = .17 vs. $M_D = .48$, SD = .20); and a main effect of status of CI, $F(1, 124) = 9.67$, MSE = .045, $p = .002$, $\eta_p^2 = .07$, ($M_{in} = .64$, SD = .22 vs. $M_{out} = .56$, SD = .22). Further, there was a significant type of item × retention interval interaction, $F(1, 124) = 7.71$, MSE = .018, $p = .006$, $\eta_p^2 = .06$, showing that in the immediate condition, there was no difference between the recognition of the AI and the recognition of the S5 ($p = .69$), whereas in the delay condition, the AI was significantly more recognised than the S5 $M_{AI} = .53$, SD = .22 vs. $M_{S5} = .42$, SD = .18, $p < .001$). There was also a significant type of item × status of CI interaction, $F(1, 124) = 4.38$, MSE = .018, $p = .04$, $\eta_p^2 = .03$, revealing that when CIs were presented within the lists, there were significantly higher levels of recognition for the AI than for the S5 ($M_{AI} = .68$, SD = .21 vs. $M_{S5} = .59$, SD = .22, $p < .001$), whereas when CIs were absent, this difference was not significant ($p = .39$). However, the type of item × retention interval × status of CI interaction was not significant, $F(1, 124) = 1.48$, MSE = .018, $p = .23$, $\eta_p^2 = .01$.

Results for the Pair TI–S5: The following results were obtained: a main effect of type of item $F(1, 124) = 12.75$, MSE = .025, $p = .001$, $\eta_p^2 = .09$, ($M_{S5} = .58$, SD = .23 vs. $M_{TI} = .51$, SD = .25); a main effect of retention interval $F(1, 124) = 92.73$, MSE = .040, $p < .001$, $\eta_p^2 = .43$, ($M_I = .66$, SD = .20 vs. $M_D = .42$, SD = .20); a main effect of status of CI $F(1, 124) = 20.23$, MSE = .04, $p < .001$, $\eta_p^2 = .14$, ($M_{in} = .60$, SD = .24 vs. $M_{out} = .49$, SD = .23); a significant type of item × retention interval interaction, $F(1, 124) = 20.99$, MSE = .025, $p < .001$, $\eta_p^2 = .15$; a significant type of item × status of CI interaction, $F(1, 124) = 19.38$, MSE = .025, $p < .001$, $\eta_p^2 = .14$; and a significant retention interval × status of CI interaction, $F(1, 124) = 9.89$, MSE = .040, $p = .002$, $\eta_p^2 = .07$.

These interactions were qualified by a significant type of item × retention interval × status of CI interaction, $F(1, 124) = 35.10$, MSE = .025, $p < .001$, $\eta_p^2 = .22$, indicating that when CIs were absent from the lists, false recognition of the TI was not different between the immediate and delay conditions ($p = .37$), whereas recognition of the S5 was significantly higher in the immediate condition than in the delay condition ($M_I = .75$, SD = .15 vs. $M_D = .38$, SD = .18, $p < .001$). When CIs were present in the lists, both items showed significantly higher levels of veridical recognition in the immediate condition than one week later ($M_I = .78$, SD = .15 vs. $M_D = .43$, SD = .18, $p < .001$, for the TI; $M_I = .74$, SD = .12 vs. $M_D = .44$, SD = .24, $p < .001$, for the S5).

Appendix 3

Results for the Pair AI–U: The following results were obtained: a main effect of type of item, $F(1, 124) = 501.60$, MSE = .027, $p < .001$, $\eta_p^2 = .80$, ($M_{AI} = .64$, SD = .23 vs. $M_U = .17$, SD = .19); a significant type of item × retention interval interaction, $F(1, 124) = 97.18$, MSE = .027, $p < .001$, $\eta_p^2 = .44$, showing that the difference in recognition between the AI and U is higher in the immediate condition ($M_{AI} = .75$, SD = .17 vs. $M_U = .09$, SD = .12, $p < .001$) than in the delay condition ($M_{AI} = .52$, SD = .22 vs. $M_U = .26$, SD = .20, $p < .001$); and a significant type of item × status of CI interaction, $F(1, 124) = 11.88$, MSE = .027, $p = .001$, $\eta_p^2 = .09$, which indicates that for the AI, veridical recognition was higher than false recognition ($M_{in} = .68$, SD = .23 vs. $M_{out} = .59$, SD = .21, $p < .001$), whereas the conditions of in/out lists of CIs did not influence the recognition of the U ($p = .10$). A significant type of item × retention interval × status of CI interaction was not found, $F(1, 124) = 2.31$, MSE = .027, $p = .13$, $\eta_p^2 = .02$.

Results for the Pair TI–U: The following results were obtained: a main effect of type of item, $F(1, 124) = 518.22$, MSE = .018, $p < .001$, $\eta_p^2 = .81$, ($M_{TI} = .54$, SD = .26 vs. $M_U = .15$, SD = .16); a main effect of status of CI, $F(1, 124) = 9.69$, MSE = .040, $p = .002$, $\eta_p^2 = .07$, ($M_{in} = .38$, SD = .21 vs. $M_{out} = .31$, SD = .19); a significant type of item × retention interval interaction, $F(1, 124) = 118.12$, MSE = .018, $p < .001$, $\eta_p^2 = .49$; a significant type of item × status of CI interaction, $F(1, 124) = 50.26$, MSE = .018, $p < .001$, $\eta_p^2 = .29$; and a significant retention interval × status of CI, $F(1, 124) = 20.10$, MSE = .040, $p < .001$, $\eta_p^2 = .14$.

These interactions were qualified by a significant type of item × retention interval × status of CI, $F(1, 124) = 29.70$, MSE = .018, $p < .001$, $\eta_p^2 = .19$, indicating that when CIs were omitted from the lists, false recognition of the TI did not differ between the immediate and delay conditions ($p = .94$), whereas the U was significantly less incorrectly recognised in the immediate condition than one week later ($M_I = .09$, SD = .12 vs. $M_D = .26$, SD = .19, $p < .001$). Additionally, when the CIs were presented, the TI was significantly more correctly recognised in the immediate condition than one week later ($M_I = .84$, SD = .13 vs. $M_D = .43$, SD = .25, $p < .001$), whereas the U was less incorrectly recognised in the immediate condition than in the delay condition ($M_I = .07$, SD = .08 vs. $M_D = .20$, SD = .15, $p < .001$, $p < .001$).