

Cognition and Neurosciences

False memories in Lewy-body disease

SALVADOR ALGARABEL,¹ ALFONSO PITARQUE,¹ ALICIA SALES,¹ JUAN CARLOS MELÉNDEZ¹ and JOAQUÍN ESCUDERO²

¹University of Valencia, Valencia, Spain

²Department of Neurology, General Hospital, Valencia, Spain

Algarabel, S., Pitarque, A., Sales, A., Meléndez, J. C. & Escudero, J. (2015). False memories in Lewy-body disease. *Scandinavian Journal of Psychology*, 56, 599–606.

Recently, de Boysson, Belleville, Phillips *et al.* (2011) found that patients with Lewy-body disease (LBD) showed significantly lower rates of false memories than healthy controls, using the Deese–Roediger–McDermott (DRM) experimental procedure. Given that this result could be explained by the practically null rate of true recognition in the LBD group (0.09), we decided to replicate the study by de Boysson *et al.* (2011), but including a new condition that would maximize the true recognition rate (and analyze its effect on the rate of false memories). Specifically, in a DRM experiment, we manipulated (within subjects) two study and recognition conditions: in the “immediate” condition, both the LBD patients and the control group of healthy older people received a different recognition test after each study list (containing twelve words associated with a non-presented critical word), while in the “delayed” condition (similar to the one in de Boysson *et al.*, 2011), the participants received the entire series of study lists and then took only one recognition test. The results showed that, in both samples, the “immediate” condition produced higher corrected rates of both true and false recognition than the “delayed” condition, although they were both lower in the LBD patients, which shows that these patients are capable of encoding and recognizing the general similitude underlying information (gist memory) in the right conditions.

Key words: False memory, Lewy-body disease, memory, DRM.

Alfonso Pitarque, Department of Methodology-Psychology, University of Valencia, Avenida Blasco Ibanez 21, 46010 Valencia, Spain.

E-mail: pitarque@uv.es

INTRODUCTION

Memory is a reconstructive process that sometimes makes errors. False memories are some of the most widely studied errors, especially using the experimental procedure by Deese–Roediger–McDermott (DRM; Roediger and McDermott, 1995). With this procedure, subjects study lists of words (e.g., wind, fresh, oxygen, etc.) that are associated with a non-studied critical lure (e.g., air), and they are more likely to falsely recall or recognize these critical lures (e.g., air) than unrelated lures (e.g., sleep). This procedure has been used to show that false memories increase with age in healthy people (e.g., Balota, Cortese, Duchek *et al.*, 1999; Dennis, Kim & Cabeza, 2007; Norman & Schacter, 1997). By contrast, several studies have shown that false recognition is significantly lower in patients with Alzheimer’s disease (AD; a disease that accounts for most cases of degenerative dementia) than in healthy older adults, but only after controlling for false alarms to unrelated lures (e.g., Balota *et al.*, 1999; Budson, Sitarski, Daffner & Schacter, 2002; Hudon, Belleville, Souchay, Gély-Nargeot, Chertkow & Gauthier, 2006). However, if we do not control the basal level of false alarms, AD patients show false recognition rates of critical lures that are similar to or greater than those found in healthy people of the same age (e.g., Balota *et al.*, 1999; Plancher, Guyard, Nicolas & Piolino, 2006; Watson, Balota & Sergent-Marsall, 2001). The reason it is important to correct for baseline false alarms in patients with dementia is that these patients tend to show a more liberal response bias than that of healthy older adults (e.g., Budson, Wolk, Chong & Waring, 2006; Snodgrass and Corwin, 1988), which increases their rates of hits and false alarms, thus reducing their corrected rates of both true and false recognition compared to healthy adults. These results have been interpreted (e.g., Balota *et al.*, 1999; Budson, Todman

& Schacter, 2006; Gallo, Shahid, Olson, Solomon, Schacter & Budson, 2006) as indicating that these AD patients seem to have a limited capacity to acquire, retain or recover the general similitude underlying the information (or gist memory; Reyna & Brainerd, 1995), a capacity that is preserved in healthy adults. For other authors (e.g., Gallo, Sullivan, Daffner, Schacter & Budson, 2004), the more liberal response bias of dementia patients is due to the fact that they tend to automatically base their responses on familiarity to make up for their episodic memory deficits (e.g., on item-specific recollection, the recollection-based monitoring process or the recall-to-reject process; e.g., Gallo *et al.*, 2004), increasing their rates of both hits and false alarms.

However, the role of false memories in other forms of dementia has received much less attention in the literature. Only the study by de Boysson *et al.* (2011) analyzed the role of false recognition in Lewy-body dementia (LBD), the second most common type of degenerative dementia in older people, accounting for 10–15% of dementia cases (see McKeith, Mintzer, Aarsland *et al.*, 2004; McKeith, Dickson, Lowe, *et al.* 2005). De Boysson *et al.* (2011) applied the DRM procedure to a sample of 10 LBD patients, 15 patients with frontotemporal dementia (FTD), and 30 balanced controls. They found that both groups of patients showed a reduced level of false recognition of the critical targets, but this reduction was greater in patients with LBD than in patients with FTD.

Nonetheless, the results from de Boysson *et al.* (2011) raise some questions: for example, if we observe the data, we realize that the LBD patients hardly show real learning of the lists (true recognition = hit rate – unrelated false alarm rate = 0.09). If there is no learning, it is difficult to talk about false recognition, given that it is based on learning the associative relationships

among the words on the study lists. Instead, their data seem to indicate a mere liberal response bias in the two groups of patients (which can be noted in their high rates of hits and false alarms). The low rate of learning in the LBD patients could be due to the fact that the study task proposed by de Boysson *et al.* (2011) might be too long for patients with dementia (as they had to study 12 study lists in a row with 12 words in each, that is, memorize 144 stimuli), which could lead to a floor effect in their rates of true recognition. Therefore, the first objective of the present study is to give the LBD patients a study task that maximizes the learning of the lists. Specifically, a different recognition test will be made for each study list containing 12 words associated with a non-presented critical lure. This condition will be called “immediate”, compared to the one proposed by de Boysson *et al.* (2011), which we will call “delayed”, where only a recognition test was applied after all study lists had been studied. With this immediate condition, we want to achieve higher true recognition rates in LBD patients than those obtained by Boysson *et al.* (2011) in their delayed condition, which will probably modify the false recognition rates.

Second, de Boysson *et al.* (2011) controlled the overall level of false alarms on the false recognition of the critical targets by means of an Analysis of Covariance (ANCOVA) (using the unrelated false alarms as covariate), and not by subtracting the

rate of unrelated false alarms from the rate of false recognition of the critical targets, which is the usual procedure (see e.g., Budson, Sitarski, Daffner & Schacter, 2002; Budson, Sullivan, Mayer *et al.*, 2002; Budson, Todman & Schacter, 2006; Hudon *et al.*, 2006). The use of the ANCOVA to estimate the false recognition of the critical targets could pose a methodological problem, as the ANCOVA requires the random assignment of the subjects to the treatments in order to justify the attribution of causality (e.g., Tabachnick & Fidell, 1989). This statistical analysis could in itself be the reason for the results mentioned above, as it is quite unexpected to achieve, from a very low rate of true recognition in the LBD group (0.09), a false memory rate that is lower than the control group, but moderately high (0.41).

Based on all of these problems, we intend to apply the DRM procedure to a sample of LBD patients (and their respective controls) in order to, first, increase the level of true recognition, and then later find out whether the LBD patients actually show the capacity to elicit false memories (correctly corrected), or whether they merely show a liberal response bias. In order to maximize the possibilities of the LBD patients really learn the study lists (thus increasing the possibilities that they will later elicit false memories), we will manipulate two within-subject study and recognition conditions. On one hand, in the *delayed* condition (see Fig. 1), we will use a similar procedure as in de

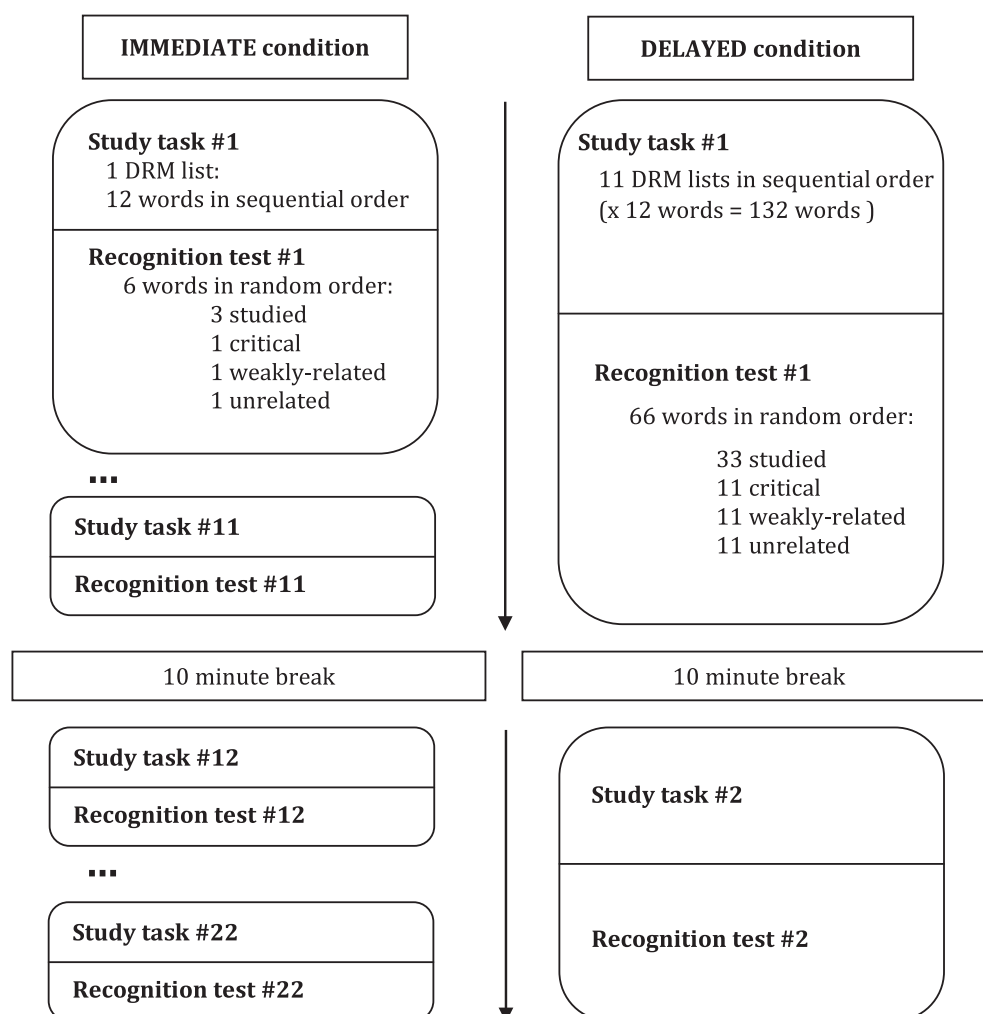


Fig. 1. Explanatory diagram of the two experimental conditions (see text for details).

Boysson *et al.* (2011), that is, sequentially presenting to each participant 11 study lists (each made up of 12 words associated with a non-studied critical lure), followed by a later recognition test containing 66 items. Half of the items will have been studied before on one of the study lists, and the other half will consist of one-third critical lures, one-third weakly-related lures (i.e., the 13th associate of each critical lure), and one-third unrelated lures. On the other hand, in the *immediate* condition (see Fig. 1) the participants will study each study list separately (each made up of 12 words associated with a non-studied critical lure). After each list, a recognition test will be performed (thus taking 11 recognition tests for the 11 lists studied). On each test, participants will have to recognize six items, three old and three new (of which, one will be the critical lure for each list, another a weakly-related lure, and the third an unrelated lure). Our objective with this immediate condition is to maximize the probability of increasing true recognition and eliciting false memories because it is well known that this type of presentation leads to greater levels of false recognition in healthy adults than presenting all the study lists together (e.g., Brainerd, Reyna & Forrest, 2002; Toglia, Neuschatz & Goodwin, 1999; Tussing & Greene, 1997).

We hypothesize that if the LBD patients show an incapacity to learn the general similitude underlying the information (gist memory), both their true recognition and false recognition rates should be the same in the immediate and delayed conditions, while if they maintain this capacity to a certain extent, then both their true recognition and false recognition rates will be greater in the immediate condition than in the delayed condition. Moreover, following this argument, the controls should also show higher rates of true and false recognition in the immediate condition than in the delayed condition. Logically, both rates should be higher than those found in the LBD sample, given that, among other reasons, patients with dementia show a greater associative-binding deficit than healthy controls (see e.g., Old & Naveh-Benjamin, 2008), or they make worse use of the recollection-based monitoring process (see e.g., Gallo *et al.*, 2004) to reduce their rates of false alarms.

METHOD

Participants

Eleven patients with a clinical diagnosis of probable Lewy-Body Disease (LBD) and 15 healthy adults serving as controls participated in the experiment. The LBD patients were recruited from the Neurology Department of the General Hospital in Valencia, Spain. Diagnosis was carried out after studying the medical history of each patient and performing a physical, neurological, neuropsychological, and psychiatric examination. Diagnosis was based on consensus criteria (McKeith *et al.*, 2004, 2005), identifying a progressive disabling mental impairment leading to dementia attentional impairments and disproportionate problem solving and visuospatial difficulties, fluctuation in cognitive function, persistent well-formed visual hallucinations, and spontaneous motor features of Parkinsonism. Exclusion criteria for the LBD group were significant past episodes of psychiatric or neuropsychological disorder, presence of other causes of dementia or brain illnesses (vascular injury, surgical brain procedure, craniocerebral trauma), a history of alcohol or drug abuse, or a very serious attention deficit that would make it impossible to perform the experimental test. The control group was comprised of healthy volunteers from centers specializing in elderly care residing in Valencia who showed no signs of depression, alcoholism, drug intake, dementia, or any other neurological disorder.

Regarding the demographic data, both groups were matched on age (mean of the LBD sample = 81.09 years, range 70–91 years; mean of the control sample = 78.07 years, range 75–85 years; $t_{12,68} = 1.53$, $p = 0.15$), gender (3 men, 8 women in the LBD sample; 2 men, 13 women in the control group; $\chi^2 = 0.79$, $p = 0.37$) and years of education (mean of the LBD sample = 6.27 years, range 0–11 years; mean of the control sample = 5.67 years, range 0–11 years; $t_{24} = 0.33$, $p = 0.75$). Furthermore, all participants underwent extensive psychometric evaluation (see neuropsychological data in Table 1). The neuropsychological battery consisted of: the Global Deterioration Scale (GDS; Reisberg, Ferris & De Leon, 1982); Mini-Mental State Examination (MMSE; Folstein, Folstein & McHugh, 1975); Geriatric Depression Scale (Sheikh & Yesavage, 1986); Functional evaluation: the Barthel index (Mahoney & Barthel, 1965); Cummings Neuropsychiatric Inventory (Cummings, Mega, Gray, Rosenberg-Thompson, Carusi & Gornbein, 1994); Verbal Fluency-Semantic, Verbal Fluency-Phonological and verbal abstraction (Peña-Casanova, 2005); Boston Naming Test (Kaplan, Goodglass & Weintraub, 1983); Logical Memory subtest of the Wechsler Memory Scale III (Wechsler, 2004); Trail Making Test-A (Reitan, 1958); Rey-Osterrieth Complex Figure (Rey, 1999); and the Stroop Test (Golden, 2005).

Table 1. Means (and SE) of the neuropsychological data

	LBD (n = 11)	Healthy controls (n = 15)	Significant differences* ($p < 0.05$)
Global Deterioration Scale	3.36 (0.20)	1.00 (0.00)	*
Mini-Mental State Examination	19.00 (1.53)	27.47 (0.51)	*
Geriatric Depression Scale	5.45 (1.14)	2.50 (0.44)	*
Barthel Index	74.09 (5.98)	98.33 (0.79)	*
Cummings Neuropsychiatric Inventory	27.82 (6.47)	0.00 (0.00)	*
Verbal fluency Phonemic	6.82 (0.97)	16.73 (1.81)	*
Verbal fluency Semantic	11.73 (2.16)	16.80 (1.98)	*
Verbal abstraction	3.09 (0.69)	5.20 (0.35)	*
Boston Naming Test	6.64 (0.75)	9.00 (0.46)	*
Logical-Memory Immediate Units	3.73 (0.91)	19.73 (2.20)	*
Logical-Memory Immediate Themes	5.18 (0.92)	12.40 (0.73)	*
Trail Making Test-A (seconds)	105.36 (29.44)	86.13 (6.85)	*
Visual Memory Test—Immediate	8.09 (3.55)	25.73 (2.53)	*
Visual Memory Test—Delayed	1.09 (0.71)	8.70 (2.01)	*
Stroop Interference (T standard scores)	50.18 (1.79)	59.33 (3.02)	*

Notes: Data represent raw scores except those specified in brackets.

*Indicates worse performance in the LBD sample.

Statistical analyses showed that the LBD group was cognitively impaired compared to healthy controls (see Table 1). However, the Verbal fluency semantic (animals) test and the Trail Making Test-A showed no significant differences between the two groups.

Materials and procedure

Forty-four study lists of 15 semantically-related words were created. Materials were selected from the lists of semantic associates published by Fernández, Díez and Alonso (2012). Each study list was composed of 12 words that were associatively related to a non-presented critical word, as is common in DRM experiments. Each list of words was presented in decreasing order of associative strength, starting with the word most closely associated with the target. On each list we reserved the last three words for use as weakly-related lures in the recognition phase. The 44 lists were divided into four blocks of 11 lists each, and the blocks were balanced on mean frequency (Block 1: frequency of 55%; Block 2: frequency of 56%; Block 3: frequency of 55%; Block 4: frequency of 55%). We manipulated two conditions with two blocks in each. In the delayed presentation condition (see Fig. 1), participants were instructed to listen carefully to the 11 word lists containing 12 words each (each list separated by a period of 5 seconds), and they were told that their memory would subsequently be tested. Words were read at a rate of about 1 s per word. The presentation of the 11 lists was followed by a recognition phase containing 66 items (half studied and half not studied; see below). Next, after a rest period of 10 minutes, participants studied a second block of 11 lists (containing 12 words each) and then performed a second recognition test with another 66 items. In the immediate presentation condition (see Fig. 1), participants were told to listen carefully to each list of 12 words, and that their memory would be tested after each list. After each study list, the participants performed a recognition test containing 6 items (half of them studied and half not studied; see below), repeating this process 11 times. Then, after a 10-minute rest period, they received a second block of 11 study lists (each containing 12 words), each followed by its corresponding 6-word recognition test. Blocks were counterbalanced across participants and across study and test conditions.

On each of the two recognition tests in the delayed condition, there were 33 studied items and 33 non-studied items (or lures). The 33 studied items consisted of three words from each of the 11 lists presented. Because the words on each list were classified according to their degree of semantic link to the critical target, we chose words in the same position on each list, drawn from input positions 3, 9, and 11. The 33 non-studied words were of three types: 11 critical lures, 11 weakly-related lures (the 13th associate of each list), and 11 unrelated lures (words that were not related semantically to any of the critical targets). On each of the 22 recognition tests in the immediate condition, there were three studied items (words selected from positions 3, 9 and 11 of each list) and three non-studied lures (the critical lure, the weakly-related lure or 13th associate of each list, and one unrelated lure). On all the recognition tests, participants were instructed to decide whether the words they were hearing consisted of studied or non-studied items from the former study list (old/new judgment). There was no time limit for responding in the recognition phase.

The patients and healthy elderly controls individually completed the neuropsychological evaluation and the two experimental conditions in two sessions (one day the delayed presentation condition and one week later the immediate presentation condition, counterbalanced across participants). The stimuli were presented in the center of a computer screen, in black lowercase Courier New Bold 18-point letters, on a white background and read aloud by the examiner. Responses were given verbally and noted by the examiner.

RESULTS

The data from the DRM tasks were analyzed by means of mixed analyses of variance with 2 groups (LBD vs controls, between subjects) \times 2 presentation conditions (delayed vs immediate,

within subjects). True recognition performance was evaluated by subtracting the proportion of false alarms to unrelated lures from the proportion of hits to studied words. False recognition was evaluated by referring to the proportion of false alarms to critical lures to the baseline of false alarms to unrelated lures. We also calculated the sensitivity (A') and response bias (B''_D) for true and false recognition (see e.g., Hudon *et al.*, 2006). As Snodgrass and Corwin (1988) point out, A' is more sensitive than d' in memory-impaired populations. The A' score ranges between 0 and 1, with 0.5 representing performance at the chance level, and higher scores indicating greater sensitivity and accuracy. The B''_D score ranges between -1 and 1, and a 0 score indicates a neutral response criterion, negative values indicate liberal or lenient responding, and positive values reflect a conservative bias.

Regarding true recognition, a 2×2 analysis of variance on hits minus unrelated false alarms ($H - FAU$; see Table 2) showed that the main effects of both the presentation conditions ($F_{1,24} = 55.39$, $p = 0.0001$, $eta^2_p = 0.70$) and groups ($F_{1,24} = 23.92$, $p = 0.0001$, $eta^2_p = 0.50$) were significant, indicating better true recognition in the control group (mean = 0.45) than in the LBD group (mean = 0.19), and in the immediate condition (mean = 0.43) than in the delayed condition (mean = 0.21). The interaction between the two variables was not significant ($F_{1,24} < 1$). Comparing the true recognition means in the LBD group to the value of 0 (which would indicate a complete absence of learning), we observed that in the immediate condition, this mean (0.30; see Table 2) differed significantly from the value of 0 ($t_{10} = 5.34$, $p = 0.0001$), while the difference between the mean of the LBD group in the delayed condition (0.08; a figure that coincides exactly with what was obtained by de Boisson *et al.*, 2011) and

Table 2. Mean proportions (and SE) of hits, false alarms on unrelated lures, false alarms on critical lures and false alarms on related lures for immediate and delayed conditions and control and LBD groups, and their estimations of true and false recognition with their Sensitivity (A') and Response Bias (B''_D) indices.

	Control group		LBD group	
	Immediate	Delayed	Immediate	Delayed
Hits (H)	0.66 (0.04)	0.61 (0.06)	0.67 (0.05)	0.58 (0.07)
False alarms on unrelated lures (FAU)	0.09 (0.04)	0.27 (0.05)	0.37 (0.05)	0.50 (0.06)
False alarms on critical lures (FAC)	0.56 (0.06)	0.63 (0.06)	0.67 (0.07)	0.63 (0.07)
False alarms on related lures (FAR)	0.21 (0.05)	0.38 (0.06)	0.45 (0.06)	0.44 (0.07)
True recognition (H - FAU)	0.56 (0.04)	0.34 (0.04)	0.30 (0.05)	0.08 (0.05)
A'	0.88 (0.02)	0.76 (0.02)	0.73 (0.02)	0.63 (0.03)
B''_D	0.59 (0.10)	0.12 (0.07)	0.03 (0.12)	-0.07 (0.09)
False recognition (FAC - FAU)	0.46 (0.05)	0.37 (0.04)	0.30 (0.06)	0.13 (0.05)
A'	0.85 (0.02)	0.77 (0.03)	0.73 (0.03)	0.67 (0.03)
B''_D	0.56 (0.12)	0.11 (0.08)	-0.08 (0.14)	-0.09 (0.09)

the value of 0 was marginally significant ($t_{10} = 2.04$, $p = 0.069$), indicating a level of true recognition that is only slightly different from chance in this group.

A similar 2×2 mixed analysis of variance (ANOVA) of the true recognition sensitivity index (A' , Table 2) showed a similar pattern of results as in the former analysis. That is, it showed significant main effects of both the presentation conditions ($F_{1,24} = 40.11$, $p = 0.0001$, $\eta^2_p = 0.63$) and groups ($F_{1,24} = 21.26$, $p = 0.0001$, $\eta^2_p = 0.47$), indicating better recognition in the control group (mean = 0.82) than in the LBD group (mean = 0.68), and in the immediate condition (mean = 0.80) than in the delayed condition (mean = 0.69). The interaction between the two variables was not significant ($F_{1,24} < 1$).

The analysis of uncorrected hits (H; see Table 2) showed a non-significant main effect of group ($F_{1,24} < 1$; as e.g., Hudon *et al.*, 2009), whereas the main effect of presentation conditions was marginally significant ($F_{1,24} = 3.84$, $p = 0.062$, $\eta^2_p = 0.14$), indicating a tendency to have more hits in the immediate condition (mean = 0.66) than in the delayed condition (mean = 0.59). The interaction between the two variables was not significant ($F_{1,24} < 1$).

Regarding the analysis of unrelated false alarms (FAU; see Table 2), the main effects of both the groups and presentation conditions were significant ($F_{1,24} = 16.03$, $p = 0.001$, $\eta^2_p = 0.40$ and $F_{1,24} = 18.07$, $p = 0.0001$, $\eta^2_p = 0.43$, respectively), indicating that the LBD group committed more unrelated false alarms (mean = 0.43) than the control group (mean = 0.18), and that more unrelated false alarms were made in the delayed condition (mean = 0.38) than in the immediate condition (mean = 0.23). The interaction between the two variables was not significant ($F_{1,24} < 1$).

Overall, the results for true recognition show that the differences in learning between the control and LBD groups are not found in their hit rates (which are similar), but rather in the fact that the LBD group makes significantly more unrelated false alarms than the control group, probably because of worse use of processes like recall-to-reject or recollection of item-specific information to reduce false alarm rates (see e.g., Abe, Fujii, Nishio *et al.*, 2011; Gallo *et al.*, 2004). Another explanation could be that the LBD group has a more liberal response bias than the control group.

Therefore, we analyzed whether the LBD group shows a more liberal response bias than the control group (which would artificially increase the rates of both hits and false alarms) by means of a 2×2 mixed ANOVA on the true recognition bias (B''_D , Table 2), which showed significant main effects of both the presentation conditions ($F_{1,24} = 14.85$, $p = 0.001$, $\eta^2_p = 0.38$) and the groups ($F_{1,24} = 10.22$, $p = 0.004$, $\eta^2_p = 0.30$), indicating a more conservative response bias in the control group (mean = 0.36) than in the LBD group (mean = -0.02), and in the immediate condition (mean = 0.31) than in the delayed condition (mean = 0.03). The interaction between the two variables was also significant ($F_{1,24} = 5.92$, $p < 0.05$, $\eta^2_p = 0.20$). Post hoc Bonferroni tests to analyze this interaction (Table 2) showed a non-significant difference between the immediate and delayed conditions in the LBD patients (means 0.03 and -0.07 , respectively; $t_{10} = 1.12$), which indicates that

these patients are equally neutral in their responses in both conditions, whereas this difference was significant in the control group (means of 0.59 and 0.12, for the immediate and delayed conditions, respectively; $t_{14} = 4.39$, $p = 0.001$), indicating that the controls are much more conservative in the immediate condition than in the delayed one. Thus, the results of this significant interaction show that the LBD patients have a neutral response bias, which confirms that the results we found for true recognition cannot be explained by a mere liberal bias in their responses.

Regarding corrected false recognition of the critical lures (FAC – FAU; Table 2), a 2×2 analysis of variance showed a significant main effect of both the presentation conditions ($F_{1,24} = 12.87$, $p = 0.001$, $\eta^2_p = 0.35$) and the groups ($F_{1,24} = 9.94$, $p = 0.004$, $\eta^2_p = 0.29$), indicating that the control group committed more false recognitions of critical lures (mean = 0.42) than the LBD group (mean = 0.21), and that the immediate condition led to more false recognitions of critical lures (mean = 0.38) than the delayed condition (mean = 0.25), as expected. The interaction between the two variables was not significant ($F_{1,24} < 1$). Comparing the false recognition means of the LBD group with the value of 0 (which would indicate a complete absence of false memories), we observed that in the delayed condition, this mean (0.13; see Table 2) differed significantly from the value of 0 ($t_{10} = 2.67$, $p < 0.05$), as did the mean of the immediate condition (0.30; $t_{10} = 4.48$, $p = 0.001$), which indicates that the LBD group is capable of eliciting false memories in both conditions. However, it should be pointed out that the mean of false recognitions of critical lures in the LBD group in the delayed condition (0.13) is significantly lower than what was observed by de Boysson *et al.* (2011, which was 0.41), which seems to suggest that ANCOVA may overestimate false recognition.

A similar 2×2 mixed ANOVA on the false recognition sensitivity index (A' , Table 2) showed a similar pattern of results to the one found for false recognition; that is, it showed significant main effects of both the presentation conditions ($F_{1,24} = 13.01$, $p = 0.001$, $\eta^2_p = 0.35$) and the groups ($F_{1,24} = 11.52$, $p = 0.002$, $\eta^2_p = 0.32$), indicating better false recognition discrimination in the control group (mean = 0.81) than in the LBD group (mean = 0.70), and in the immediate condition (mean = 0.79) than in the delayed condition (mean = 0.72). The interaction between the two variables was not significant ($F_{1,24} < 1$).

Regarding the analysis of the false recognition bias (B''_D , Table 2), the 2×2 mixed ANOVA showed significant main effects of both the presentation conditions ($F_{1,24} = 6.32$, $p < 0.05$, $\eta^2_p = 0.21$) and groups ($F_{1,24} = 11.26$, $p = 0.003$, $\eta^2_p = 0.32$), indicating a more conservative bias in the control group (mean = 0.34) than in the LBD group (mean = -0.08), and in the immediate condition (mean = 0.24) than in the delayed condition (mean = 0.01). The interaction between the two variables was also significant ($F_{1,24} = 5.66$, $p < 0.05$, $\eta^2_p = 0.19$). Post hoc Bonferroni tests to analyze this interaction (Table 2) showed a non-significant difference between the immediate and delayed conditions in the LBD patients (means -0.08 and -0.09 , respectively; $t_{10} < 1$), which clearly indicates that the LBD group uses a neutral response criterion, whereas this difference was

significant in the control group (means of 0.56 and 0.11, for the immediate and delayed conditions, respectively; $t_{14} = 3.85$, $p = 0.002$), exactly replicating the results found for true recognition bias.

Finally, regarding the analysis of false recognitions of corrected weakly-related lures (FAR – FAU; see Table 2), the main effect of the variable group was significant ($F_{1,24} = 5.95$, $p < 0.05$, $\eta^2_p = 0.20$), indicating that the control group committed more false recognitions on related lures (mean = 0.12) than the LBD group (mean = 0.01). The main effect of the presentation conditions was marginally significant ($F_{1,24} = 3.78$, $p = 0.064$, $\eta^2_p = 0.14$), indicating a tendency to commit more false recognitions on related lures in the immediate condition (mean = 0.10) than in the delayed condition (mean = 0.03). The interaction between the two variables was also marginally significant ($F_{1,24} = 3.99$, $p = 0.057$, $\eta^2_p = 0.14$), probably due to a floor effect in the LBD group in the *delayed* condition (mean = -0.06). All the means for false alarms on related lures were located between the means for false alarms on critical lures and the means for false alarms on unrelated lures (see Table 2), supporting our experimental procedure (see Roediger & McDermott, 1995).

DISCUSSION

Our results, in agreement with the hypotheses proposed, show that by maximizing the learning of the lists by employing an immediate condition, we were able to increase the false recognition of critical lures in LBD patients (as well as in healthy ones). In other words, this LBD group is capable of eliciting false memories in the immediate condition and, to a lesser degree, in the delayed condition, revealing that LBD patients show a certain capacity to learn the general similitude underlying the information (gist memory). However, this capacity is lower than that of a control group of healthy older people, which places the performance of the LBD group on DRM tasks at a level comparable to the performance of AD patients (Balota *et al.*, 1999; Budson *et al.*, 2002; Hudon *et al.*, 2006). These results also coincide with those observed when comparing LBD and AD patients on cognitive or neuropsychiatric measures (e.g., Walker, McKeith, Rodda *et al.*, 2012). However, even though the rate of true recognition in our data in the delayed condition is exactly the same as what de Boysson *et al.* (2011) obtained, our rate of false recognition of critical lures in this condition is significantly lower than what was obtained by these authors, which seems to indicate that controlling the basal rate of false alarms to unrelated lures using ANCOVA does not appear to be a correct statistical procedure for estimating the false recognition of critical lures.

Our results also demonstrate that the corrected rate of false recognition tends to be proportional to the corrected rate of true recognition (see Table 2; results similar to those obtained e.g., by Hudon *et al.*, 2006), in both LBD patients and controls, and in both the delayed and immediate conditions, showing that false recognition is based on learning the associative relationships between the words on the study lists and that this learning must be situated at intermediate levels to optimally elicit false memories (if the true recognition reaches a ceiling or floor effect, there cannot be any elicitation of false memories). Only in the

immediate condition in the control group did true recognition significantly surpass false recognition (means of 0.56 and 0.46, respectively; $t_{14} = 2.45$, $p < 0.05$; see Table 2), which would indicate that healthy older people are using another mechanism, in addition to gist memory, to respond. This idea becomes reinforced if we consider that the control group uses a much more conservative response bias in the immediate condition than in the delayed condition, indicating that in the immediate condition they use a response mechanism that allows them to significantly reduce their rate of false alarms. This mechanism could be, for example, a conscious mechanism such as item-specific recollection or recall-to-reject, through which healthy controls are able to minimize their rate of false alarms to unrelated lures, a mechanism that LBD patients cannot use, due to their episodic memory deficits.

De Boysson *et al.* (2011) suggest that patients with frontal lobe lesions (and AD and LBD patients, frontotemporal dementia patients, etc.) show both automatic inhibitory and executive control deficits that keep them, for example, from ignoring irrelevant information or inhibiting their tendency to respond “old” to critical lures on the basis of familiarity alone (or gist memory). Healthy people, on the other hand, can use their item-specific recollection capacity to suppress this tendency (based on familiarity alone) to say “yes”. Supporting this idea, Budson, Sullivan, Mayer *et al.* (2002) found that the repetition of the words on the study task produced a reduction in false recognition in healthy people (improving their item-specific recollection capacity with practice), while AD patients maintain a stable rate of false memories across repetitions (showing their inability to improve their item-specific recollection, due to responding only on the basis of familiarity or gist memory). In our data, both the worse performance of the LBD patients on the Stroop test (which measures the ability to inhibit the irrelevant information) and the significant correlation found between Stroop scores and false memories rates in the immediate condition ($r = -0.39$, $p < 0.05$, but not in the delayed condition) seem to support this idea, although we must be cautious in that asseveration due to its correlational nature and the small sample size. The role that inhibitory and executive deficits play in both healthy and pathological aging has been demonstrated through different experimental methodologies such as retrieval-induced forgetting (e.g., Gómez-Ariza, Pelegrina, Lechuga, Suárez & Bajo, 2009) or think/no think (Anderson, Reinholz, Kuhl & Mayr, 2011), and they deserve a more detailed analysis within the DRM paradigm.

Overall, our results could be best explained by the gist memory theory (also called fuzzy trace-theory; Reyna & Brainerd, 1995), which explains true recognition as being due to the sum of two components: the successful recollection of item-specific information (a capacity that is reduced in patients with dementias like LBD, DCL, AD, etc., compared to healthy subjects) and gist memory (a capacity that is somewhat impaired, but not annulled, in patients with dementia compared to healthy patients), while false recognition would only be explained by gist memory; that is, false recognition depends on remembering gist, but not item-specific information. This theory would explain why our control group obtains higher rates on both true recognition and false recognition than the LBD group (due to their greater capacity for recollection of item-specific information and better gist memory

in the former case, and their better gist memory in the latter), or why in the immediate condition we obtain higher rates on both true recognition and false recognition than in the delayed condition (given that in the latter condition, due to the greater number of items to process on the study list, there is a reduction in both the capacity for recollection of item-specific information and the formation of a clear gist memory), or why true recognition significantly surpasses false recognition only in the immediate condition in the control group (because only here is the recollection of item-specific information effective in reducing their rates of false alarms). In the same way, our results could be explained by the activation/monitoring framework (e.g., Roediger, Watson, McDermott & Gallo, 2001), which also explains both true recognition and false recognition in terms of two components: the automatic activation of the studied words that spreads to the non-studied related lures (affecting false recognition) and a recollection-based monitoring process of conscious decision-making, whose result can be correct (affecting true recognition) or incorrect (affecting false recognition). Therefore, overall our results would support the theoretical postulates of the so-called “dual” or “two-processes memory” models (see e.g., Koen & Yonelinas, 2014; Schoemaker, Gauthier & Pruessner, 2014; Yonelinas & Jacoby, 2012, for reviews), which propose that two processes intervene in recovering (correct or incorrect) information from our memory, a conscious process of recovering episodic traces and another automatic process (called different names by different authors: recollection and familiarity, item-specific information and gist memory, explicit and implicit memory, respectively), an idea that has received considerable experimental support, even in animal experiments (e.g., Basile & Hampton, 2013). These two processes seem to rest on different neuroanatomical bases within the medial temporal lobe: recollection and episodic memory seem to be mainly related to hippocampal functioning, while familiarity is associated with perirhinal and entorhinal cortex functioning (e.g., Yonelinas, Aly, Wang & Koen, 2010; Yonelinas & Jacoby, 2012).

In summary, our study shows that LBD patients, after adequate correction of the basal level of false alarms to non-related lures, are capable of showing rates (significantly different from 0) of true recognition and false recognition (see also de Boysson *et al.*, 2011), which would indicate that they are capable, to a certain degree, of acquiring, retaining and recovering the general similitude underlying the information (or gist memory; Reyna & Brainerd, 1995).

This research was supported by Grant PSI2013-43533-R from the Dirección General de Investigación Científica y Técnica (proyectos I+D+i) of the Spanish Ministerio de Economía y Competitividad.

REFERENCES

- Abe, N., Fujii, T., Nishio, Y., Iizuka, O., Kanno, S., Kikuchi, H. *et al.* (2011). False item recognition in patients with Alzheimer's disease. *Neuropsychologia*, *49*, 1897–1902.
- Anderson, M. C., Reinholz, J., Kuhl, A. & Mayr, U. (2011). Intentional suppression of unwanted memories grows more difficult as we age. *Psychology and Aging*, *26*, 397–405.
- Balota, D. A., Cortese, M. J., Duchek, J. M., Adams, D., Roediger, H. L., III, McDermott, K. B. *et al.* (1999). Veridical and false memories in healthy older adults and in dementia of the Alzheimer's type. *Cognitive Neuropsychology*, *16*, 361–384.
- Basile, B. M. & Hampton, R. R. (2013). Recognition errors suggest fast familiarity and slow recollection in rhesus monkeys. *Learning & Memory*, *20*, 431–437.
- Brainerd, C. J., Reyna, V. F. & Forrest, T. J. (2002). Are young children susceptible to the false-memory illusion? *Child Development*, *73*, 1363–1377.
- Budson, A. E., Sitarski, J., Daffner, K. R. & Schacter, D. L. (2002). False recognition of pictures versus words in Alzheimer's disease: The distinctiveness heuristic. *Neuropsychology*, *16*, 163–173.
- Budson, A. E., Sullivan, A. L., Mayer, E., Daffner, K. R., Black, P. M. & Schacter, D. L. (2002). Suppression of false recognition in Alzheimer's disease and patients with frontal lobe lesions. *Brain*, *125*, 2750–2765.
- Budson, A. E., Todman, R. W. & Schacter, D. L. (2006). Gist memory in Alzheimer's disease: Evidence from categorized pictures. *Neuropsychology*, *20*, 113–122.
- Budson, A. E., Wolk, D. A., Chong, H. & Waring, J. D. (2006). Episodic memory in Alzheimer's disease: Separating response bias from discrimination. *Neuropsychologia*, *44*, 2222–2232.
- Cummings, J. L., Mega, M., Gray, K., Rosenberg-Thompson, S., Carusi, D. A. & Gornbein, J. (1994). The Neuropsychiatric Inventory. *Neurology*, *44*, 2308–2314.
- de Boysson, C., Belleville, S., Phillips, N. A., Johns, E. K., Goupil, D., Souchay, C. *et al.* (2011). False recognition in Lewy-body disease and frontotemporal dementia. *Brain and Cognition*, *75*, 111–118.
- Dennis, N. A., Kim, H. & Cabeza, R. (2007). Effects of aging on true and false memory formation: An fMRI study. *Neuropsychologia*, *45*, 3157–3166.
- Fernández, A., Díez, E. & Alonso, M. A. (2012). *Normas de falso recuerdo y reconocimiento [Norms of false recall and recognition]*. Retrieved 29 April 2015 from http://inico.usal.es/usuarios/gimc/normas/index_fr.asp.
- Folstein, M., Folstein, S. & McHugh, P. (1975). Mini Mental State. A practical method for grading the cognitive state of patients for the clinical. *Journal of Psychiatric Research*, *12*, 189–198.
- Gallo, D. A., Sullivan, A. L., Daffner, K. R., Schacter, D. L. & Budson, A. E. (2004). Associative recognition in Alzheimer's disease: Evidence for impaired recall-to-reject. *Neuropsychology*, *18*, 556–563.
- Gallo, D. A., Shahid, K. R., Olson, M. A., Solomon, T. M., Schacter, D. L. & Budson, A. E. (2006). Overdependence on degraded gist memory in Alzheimer's disease. *Neuropsychology*, *20*, 625–632.
- Golden, C. J. (2005). *Stroop color and word test. A manual for clinical and experimental uses*. Madrid: TEA.
- Gómez-Ariza, C. J., Pelegrina, S., Lechuga, M. T., Suárez, A. & Bajo, M. T. (2009). Inhibition and retrieval of facts in young and older adults. *Experimental Aging Research*, *35*, 83–97.
- Hudon, C., Belleville, S., Souchay, C., Gély-Nargeot, M. C., Chertkow, H. & Gauthier, S. (2006). Memory for gist and detail information in Alzheimer's disease and mild cognitive impairment. *Neuropsychology*, *20*, 566–577.
- Kaplan, E. F., Goodglass, H. & Weintraub, S. (1983). *Boston Naming Test*. Philadelphia, PA: Lea & Febiger.
- Koen, J. D. & Yonelinas, A. P. (2014). The effects of healthy aging, amnesic mild cognitive impairment, and Alzheimer's disease on recollection and familiarity: A Meta-analytic Review. *Neuropsychology Review*, *24*, 332–354.
- Mahoney, F. I. & Barthel, D. W. (1965). Functional evaluation: The Barthel index. *Maryland State Medical Journal*, *14*, 61–65.
- McKeith, I., Dickson, D. W., Lowe, J., Emre, M., O'Brien, J. T., Feldman, H. *et al.* (2005). Diagnosis and management of dementia with Lewy bodies: Third report of the DLB Consortium. *Neurology*, *65*, 1863–1872.
- McKeith, I., Mintzer, J., Aarsland, D., Burn, D., Chiu, H., Cohen-Mansfield, J. *et al.* (2004). Dementia with Lewy bodies. *Lancet Neurology*, *3*, 19–28.

- Norman, K. A. & Schacter, D. L. (1997). False recognition in younger and older adults: Exploring the characteristics of illusory memories. *Memory and Cognition*, 25, 838–848.
- Old, S. R. & Naveh-Benjamin, M. (2008). Differential effects of age on item and associative measures of memory: A meta-analysis. *Psychology and Aging*, 23, 104–118.
- Peña-Casanova, J. (2005). *Test Barcelona revisado [Revised Barcelona Test]*. Barcelona: Masson.
- Plancher, G., Guyard, A., Nicolas, S. & Piolino, P. (2009). Mechanisms underlying the production of false memories for famous people's names in aging and Alzheimer's disease. *Neuropsychologia*, 47, 2527–2536.
- Reisberg, B., Ferris, S. H. & De Leon, M. J. (1982). The global Deterioration Rating Scale: An instrument on the assessment of primary degenerative dementia. *Psychiatry*, 139, 1936–1939.
- Reitan, R. M. (1958). Validity of the trail making test as an indicator of organic brain damage. *Perceptual and Motor Skills*, 8, 271–276.
- Rey, A. (1999). *Test de copia y reproducción de memoria de figuras geométricas complejas [Copy and Reproduction of Complex Geometric Figures from Memory Test]*. Madrid: TEA editions.
- Reyna, V. F. & Brainerd, C. J. (1995). Fuzzy-trace theory: An interim synthesis. *Learning and Individual Differences*, 7, 1–75.
- Roediger, H. L., III & McDermott, K. B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21, 803–814.
- Roediger, H. L., III, Watson, J. M., McDermott, K. B. & Gallo, D. A. (2001). Factors that determine false recall: A multiple regression analysis. *Psychonomic Bulletin and Review*, 8, 385–407.
- Schoemaker, D., Gauthier, S. & Pruessner, J. C. (2014). Recollection and familiarity in aging individuals with mild cognitive impairment and Alzheimer's disease: A literature review. *Neuropsychology Review*, 24, 313–331.
- Sheikh, J. L. & Yesavage, J. A. (1986). Geriatric Depression Scale (GDS). Recent evidence and development of a shorter version). *Clinical Gerontology*, 5, 165–172.
- Snodgrass, J. G. & Corwin, J. (1988). Pragmatics of measuring recognition memory: Applications to dementia and amnesia. *Journal of Experimental Psychology: General*, 117, 34–50.
- Tabachnick, B. G. & Fidell, L. S. (1989). *Using multivariate statistics*. New York: HarperCollins.
- Toglia, M. P., Neuschatz, J. S. & Goodwin, K. A. (1999). Recall accuracy and illusory memories: When more is less. *Memory*, 7, 233–256.
- Tussing, A. A. & Greene, R. L. (1997). False recognition of associates: How robust is the effect? *Psychonomic Bulletin & Review*, 4, 572–576.
- Walker, Z., McKeith, I., Rodda, J., et al. (2012). Comparison of cognitive decline between dementia with Lewy bodies and Alzheimer's disease: A cohort study. *BMJ Open*, 2, 1–7. doi:10.1136/bmjopen-2011-000380.
- Watson, J. M., Balota, D. A. & Sergent-Marshall, S. D. (2001). Semantic, phonological, and hybrid veridical and false memories in healthy older adults and in individuals with dementia of the Alzheimer type. *Neuropsychology*, 15, 254–268.
- Wechsler, D. (2004). *Wechsler Memory Scale III*. Madrid: TEA ediciones.
- Yonelinas, A. P. & Jacoby, L. L. (2012). The process-dissociation approach two decades later: Convergence, boundary conditions, and new directions. *Memory & Cognition*, 40, 663–680.
- Yonelinas, A. P., Aly, M., Wang, W. C. & Koen, J. D. (2010). Recollection and familiarity: Examining controversial assumptions and new directions. *Hippocampus*, 20, 1178–1194.

Received 10 March 2015, accepted 19 June 2015