 DOI: 10.1111/sjop.12168

Development and Aging

Repetition increases false recollection in older people

ALFONSO PITARQUE, ALICIA SALES, JUAN CARLOS MELÉNDEZ and SALVADOR ALGARABEL

Department of Psychology-Methodology, University of Valencia, Avda Blasco Ibanez, 21, 46010 Valencia, Spain

Pitarque, A., Sales, A., Meléndez, J. C. & Algarabel, S. (2015). Repetition increases false recollection in older people. *Scandinavian Journal of Psychology, 56*, 38–44.

Aging is accompanied by an increase in false alarms on recognition tasks, and these false alarms increase with repetition in older people (but not in young people). Traditionally, this increase was thought to be due to a greater use of familiarity in older people, but it was recently pointed out that false alarms also have a clear recollection component in these people. The main objective of our study is to analyze whether the expected increase in the rate of false alarms in older people due to stimulus repetition is produced by an inadequate use of familiarity, recollection, or both processes. To do so, we carried out an associative recognition experiment using pairs of words and pairs of images (faces associated with everyday contexts), in which we analyzed whether the repetition of some of the pairs increases the rate of false alarms in older people (compared to what was found in a sample of young people), and whether this increase is due to familiarity or recollection (using a remember-know paradigm). Our results show that the increase in false alarms in older people due to repetition is produced by false recollection, calling into question both dual and single-process models of recognition. Also, older people falsely recollect details of never studied stimuli, a clear case of perceptual illusions. These results are better explained in terms of source-monitoring errors, mediated by people's retrieval expectations.

Key words: Aging, memory, associative recognition, repetition, familiarity, recollection, false recollection, older people.

Alfonso Pitarque, *Department of Psychology-Methodology*, *University of Valencia*, *Avda Blasco Ibanez*, *21*, *46010 Valencia*, *Spain*.
E-mail: pitarque@uv.es

INTRODUCTION

This research analyzes why older people commit more false alarms (FA) in recognition tasks than young people, as repetitions of the stimuli increase (e.g., Buchler, Faunce, Light, Gottfredson & Reder, 2011; Light, Patterson, Chung & Healy, 2004). A sample of older people and a sample of young people performed two associative recognition tasks in which we manipulated (within subjects) the type of stimuli (pairs of words or pairs of images) and the number of presentations of the stimuli during the study phase (once vs. twice). Remember-know judgments on the recognition responses will allow to analyze and discuss the nature of these FA.

There is considerable consensus in the literature to accept that the recognition of an item (or an association of items) can be based either on consciously eliciting contextual details associated with that item (*recollection*) or merely on an automatic sense of familiarity in the absence of contextual details (*familiarity*). Recollection is conceived of as a conscious, attention-demanding process of threshold-type qualitative retrieval, while familiarity would be an unconscious and automatic retrieval process based on the quantitative estimation of the strength of the memory trace, based on signal detection theory (SDT). Therefore, these models are known as dual models (e.g., Yonelinas, 2002; Yonelinas & Jacoby, 2012, for reviews), which have received considerable experimental support, even in animal experiments (e.g., Basile & Hampton, 2013). In contrast to the dual models, the single-process or global strength theories (e.g., Wixted, 2007; Wixted & Mickes, 2010), propose that information retrieval from the memory is based only on a quantitative estimation of the trace strength based on the SDT: what dual models call familiarity would refer to weak memories, while recollection would refer to strong memories.

Although there are several experimental procedures to estimate recollection and familiarity (e.g., Yonelinas, 2002, for a review), the remember-know (RK) procedure (Tulving, 1985) is the most widely used (Migo, Mayes & Montaldi, 2012). In the RK paradigm participants are required to report whether they recognize an item on the basis of recollection of episodic information about the study event (remembering) or the item is familiar in the absence of recollection (knowing). Even though the RK procedure has been criticized for its introspective nature, it is true that R and K responses produce results that agree with results from other methods for estimating recollection and familiarity (e.g., McCabe, Geraci, Boman, Sensenig & Rhodes, 2011; Migo *et al*., 2012), thus we consider it can be a suitable method for the study of the nature of FA.

One trait of human memory is its susceptibility to distortions and illusions (e.g., Koriat, Goldsmith & Pansky, 2000, for a review). On recognition tasks, these memory failures become evident through the analysis of FA. FA (as well as false memories, illusions, etc.) are known to increase with age (e.g., McCabe, Roediger, McDaniel & Balota, 2009; Norman & Schacter, 1997; Rhodes, Castel & Jacoby, 2008), especially in the case of events that share perceptual or conceptual characteristics (Norman & Schacter, 1997; Schacter, Koutstaal & Norman, 1997). This increase in FA is generally thought to be due to the (incorrect) use of familiarity to compensate for the episodic deficit that accompanies age (e.g., Buchler *et al*., 2011; Rhodes *et al*., 2008). However, in a meta-analytic review using the RK paradigm, McCabe *et al*. (2009) recently showed that age-associated deficits are observed through an increase in false alarms on recollection type judgments, which would indicate that these deficits are mainly due to an incorrect recollection of episodic details associated with other items studied before (source misattributions), rather than an incorrect use of familiarity (see also Dodson, Bawa & Krueger, 2007; Gallo & Roediger, 2003; Norman & Schacter, 1997), which clearly contradicts the theoretical proposals of the dual models, as it is not possible to recollect what has not been previously studied.

The associative recognition paradigm is an appropriate tool for analyzing the role that recollection and familiarity play in both correct and false recognition, as this procedure produces higher rates of FA in older people than those obtained using the item recognition procedure (Rhodes *et al*., 2008). In the associative recognition paradigm, the participants study non-related pairs of stimuli, having to discriminate on the later recognition task between intact pairs (that is, recognize pairs of stimuli presented exactly the way they were studied), rearranged pairs (or pairs of stimuli that both appeared in the study phase, but are not presented in the same pairs in the recognition phase) or new pairs (pairs of stimuli not studied before). Using this paradigm, older people have shown less ability to discriminate between intact and rearranged pairs than young people, clearly indicating that they show an associative-binding deficit (Craik & Rose, 2012; Old & Naveh-Benjamin, 2008). Furthermore, the older people produced more FA on the rearranged pairs than the young people did, and this difference increased with the number of repetitions of the stimuli during the study phase (Buchler *et al*., 2011; Light *et al*., 2004; Rhodes *et al*., 2008), which would indicate that their recognition depends more on familiarity-based processes when faced with a deficit in recollection. However, this result contradicts the results from the McCabe *et al*. (2009) meta-analysis, which show that the increase in FA in older people is mainly due to recollection-type judgments and not to familiarity.

The main objective of our study is to shed light on this contradiction in the literature (that is, to analyze whether the greater rate of FA that would be expected in elderly people due to stimulus repetition has its origins in recollection or familiarity), using an associative recognition paradigm with *RK* judgments. In the study phase of our experiment, the participants will study pairs of words or pairs of photographs (faces associated with contexts), half of which will be presented only once and the other half twice. On the later associative recognition task, a third of the pairs will be intact (half of them studied once, and the other half twice), a third will be rearranged (half of them studied once, and the other half twice), and the remaining third will be new pairs. The participant has to choose one of the these four response options: a) recognizes the pair as studied before because he/she remembers details from it (*remember* or *R* judgment); b) recognizes the pair presented as studied before because it is familiar to him/her, but without details (*know* or *K* judgment); c) recognizes the pair presented as studied before because he/she knows that the stimuli appeared before, but are now *rearranged*; d) recognizes the pair presented as not studied before because he/she knows the stimuli are *new*. That is, we applied the *RK* associative recognition paradigm in only one step (and not in two steps as is usually done, following the recommendations of Migo *et al*., 2012, trying to avoid the retrospective *RK* judgment involved after the first *yes* decision).

The idea of using materials consisting of photographs of faces associated with natural contexts is not new (e.g., Chen & Naveh-Benjamin, 2012; Kilb & Naveh-Benjamin, 2011). Various arguments can be made to justify it: images are said to have greater ecological validity than verbal materials (Weinstein & Shanks, 2010); they promote a deeper encoding of the materials (Craik & Rose, 2012); and, finally, they are better recalled than words (picture superiority effect) by both young and older people (e.g., Defeyter, Russo & McPartlin, 2009; Paivio, 1971). Moreover, this effect increases with age (Defeyter *et al*., 2009). Likewise, images are known to provoke more FA in healthy older people than in young people (e.g., Vannuci, Mazzoni, Marchetti & Lavezzini, 2012), so that they fit the objectives of the present study perfectly.

Based on the above, we hypothesize that regarding false recognition, elderly people will commit more FA than the young people, and the FA will increase with the repetitions, but only in the group of older people, and not in young people (e.g., Buchler *et al*., 2011). As young people do not show episodic impairment, they will use the repetitions to improve their strategy of “recall to reject” (e.g., Jones & Jacoby, 2001; Rotello & Heit, 2000), thus minimizing the rate of FA. However, and this is the true original objective of our study, we do not know whether this expected increase in FA in elderly people with repetitions will be due to recollection errors (such as misattributions, misrecollections, source monitoring errors, etc.; e.g., McCabe *et al*., 2009) or to errors in familiarity (due, for example, to the greater activation of the items repeated during the study; e.g., Buchler *et al*., 2011; Rhodes *et al*., 2008).

METHOD

Participants

Participants consisted of 28 young adults (Psychology undergraduates at the University of Valencia; 6 men, 22 women, mean age = 21.46 years, SD = 4.05 years) and 24 older adults (recruited from a University of Valencia course for elderly people; 6 men, 18 women, mean age = 66.63 years, SD = 5.24 years). All participants reported being in good physical and mental health, with no known memory impairments. In this regard, the mean in the Mini-Mental State Examination (Folstein, Folstein & McHugh, 1975) for the older adults was 28.62 (SD = 1.98), thus showing no memory impairment. The two groups were matched on gender ($\chi_1^2 = 0.09$), and they differed on WAIS vocabulary (Wechsler, 2001; $t_{50} = 4.42$, $p < 0.0001$), with the elderly people having more vocabulary (mean = 10.54, SD = 2.73) than the younger people (mean = 7.64, SD = 1.96), and years of education ($t_{50} = 3.60$, $p = 0.001$, with the young people having more education than the older people, means of 16.07 and 13.33 years, and SD of 0.93 and 2.91, respectively), which are a typical pattern of results (e.g., Buchler *et al*., 2011).

Materials

Word pairs. The materials consisted of 128 Spanish words with between three and nine letters selected from the Alameda and Cuetos (1995) database. The words had a mean frequency of 50.44 per two million (SD = 108.66), and a mean length of 5.30 letters (SD = 1.01). With these words, we then created four different lists of 64 pairs of words each, randomly paired: 10 randomly chosen pairs on each list were presented to the subjects once during the study phase (later making up the *intact non-repeated condition* for recognition); another 10 randomly chosen pairs were presented twice during the study phase (later making up the *repeated intact condition* for recognition); another 10 randomly chosen pairs were presented once during the study phase (and later made up the *non-repeated rearranged condition* for recognition, randomly re-pairing these stimuli in a different order from the one studied); another randomly chosen 10 pairs were presented twice during the study phase (making up the *repeated rearranged condition* for recognition, randomly re-pairing these stimuli in a different order from the one studied); another randomly chosen 20 pairs from each list were not presented during the study phase, but were presented on the recognition task as stimuli in the *new condition*. Finally, the four remaining pairs on each list were used as distractors, with two presented at the beginning and two at the end of each study task, and not being tested later. These 4 lists were counterbalanced between subjects; that is, each subject received one of these lists. Two analyses of variance on the rates of hits and FA produced by these lists confirmed that there were no significant differences between them on either of these two dependent variables (F < 1, in both cases).

Picture pairs. The materials consisted of 64 ID-card sized color photographs (145 × 160 pixels) of anonymous faces (16 of young men, 16 of young women, 16 of older men and 16 of older women) and 64 color photographs of unknown everyday scenery (800 × 600 pixels), of which one-third were open natural scenes (beaches, fields, etc.), another third were open urban scenes (streets, building, etc.), and the remaining third were interiors of building (living rooms, bedrooms, etc.). With these photographs, we then created four different lists of 64 pairs of photographs each, putting a randomly chosen face in the center of a randomly chosen scene, thus setting up the same study and recognition conditions as those used with the pairs of words. These four lists of images were also counter-balanced between subjects, and later analyses of the hits and false alarms produced by these lists confirmed that there were no significant differences between them (F < 1, in both cases).

Procedure

For about 60 minutes, the participants performed two associative recognition tasks (of pairs of words and pairs of images, counter-balanced between subjects) separated by a 30-minute period in which they took the WAIS vocabulary test (Wechsler, 2001) and, only the older people took the Mini-Mental State Examination (Folstein *et al*., 1975).

Each of the associative recognition tasks consisted of a study phase with pairs of stimuli and a recognition task with these pairs. In the study phase, 60 pairs of stimuli ± 2 distractors were presented sequentially and randomly at the center of a computer screen for 2.5 seconds each (followed by 1 second delay before the next pair). Of these 60 pairs, 10 pairs were presented once in the subjects' study phase (later making up the *non-repeated intact recognition condition*); 10 pairs were presented twice during the study phase (later making up the *repeated intact recognition condition*); 10 pairs were presented once during the study phase (later making up the *non-repeated rearranged recognition condition*, randomly re-matching these stimuli in a different order from the one studied); another 10 randomly-chosen pairs were presented twice during the study phase (making up the *repeated rearranged recognition condition*, randomly rearranging these stimuli in a different order from the one studied). The words were always presented in black 18-point Courier font on a white background.

In the recognition phase, 60 pairs of stimuli were presented sequentially and randomly at the center of a computer screen, and they remained on the screen until the subject responded to the question about whether these words or images had appeared together or not in the previous study task (depending on the recognition task they were doing). To respond, they had to choose one of these four response options:

1. yes, because I *remember* some details;
2. yes, because I *know* that they went together, but not details;
3. no, both stimuli appeared before, but they are *rearranged*;
4. no, neither stimulus appeared before, they are *new*.

In other words, participants were asked to respond "old" only to intact pairs. The text for the response options appeared at the bottom of the screen, together with the appropriate response key for each option. Of the 60 pairs of stimuli presented on the recognition task, 10 corresponded to the *non-repeated intact condition*, 10 corresponded to the *repeated intact condition*, 10 corresponded to the *non-repeated rearranged condition*, 10 corresponded to the *repeated rearranged condition*, and 20 corresponded to the *new condition*. Prior to performing the recognition task, the differences between "remembering" and "knowing" were explained to subjects using Rajaram's (1993) instructions, which, in short, make the distinction between recognition based on retrieving episodic traces from the items and recognition based on a mere sense of familiarity without any specific, episodic information. A practice study and recognition task was performed to make sure all the subjects understood the instructions, in which each subject had to call out the difference between remembering and knowing (see Migo *et al*., 2012).

Design and statistical analysis

We used an experimental design with three independent variables: 2 age groups (young vs. elderly people; between subjects) × 2 stimuli (words vs. pictures; within subjects) × 3 repetition conditions (non-repeated pairs, repeated pairs and new stimuli; within subjects). The dependent variables were the rates of correct answers (hits), FA, and the non-parametric sensitivity index (A′). The data were analyzed by means of analysis of variance. The significance level for all statistical tests was $p < 0.05$.

RESULTS

General recognition performance

We began by analyzing the global recognition data (shown in Table 1). With regard to hits (hits in remember judgments + hits in know judgments), a mixed analysis of variance (ANOVA) of 2 age groups (young vs. elderly people; between subjects) × 2 stimuli (words vs. pictures; within subjects) × 2 repetition conditions (non-repeated vs. repeated stimuli; within subjects) showed that the main effects of the three variables were significant: groups ($F_{1,50} = 27.22$, $p < 0.0001$, $\eta^2_p = 0.35$), stimuli ($F_{1,50} = 11.87$, $p = 0.001$, $\eta^2_p = 0.19$), and repetition conditions ($F_{1,50} = 95.13$, $p < 0.0001$, $\eta^2_p = 0.66$), indicating that young people hit more than elderly people (means of 0.82 and 0.64, respectively), pictures were better recognized than words (means of 0.78 and 0.68, respectively), and repeated stimuli were better recognized than unrepeated stimuli (means of 0.82 and 0.64, respectively). These results coincide with those found in the literature reviewed in the Introduction, and they support our procedure. Moreover, the stimuli × groups interaction was also significant ($F_{1,50} = 5.65$, $p = 0.02$, $\eta^2_p = 0.10$), indicating that the difference between pictures and words was greater in older people (means 0.73 and 0.55, respectively) than in young people (means 0.84 and 0.81, respectively), and supporting the idea that images are more appropriate for older people than verbal material (e.g., Defeyter *et al*., 2009).

Regarding false alarms (FA in remember judgments + FA in know judgments; see Table 1), a mixed ANOVA of 2 groups × 2 stimuli × 3 repetition conditions (non-repeated, repeated and new stimuli; within subjects) showed significant main effects for both the groups variable ($F_{1,50} = 48.30$, $p < 0.0001$, $\eta^2_p = 0.49$), indicating that elderly people committed more FA than young people (means of 0.22 and 0.07, respectively), and the repetition

Table 1. *Means (and SD) of hits, false alarms, and sensitivity index of the global results of the associative recognition tasks as a function of age groups, stimuli, and repetition conditions. Significant differences between age groups (p < 0.05) marked with **

			Older (n = 24)	Young (n = 28)
Hits	Words	Non rep	0.45 (0.24)	0.71 (0.22)*
		Rep	0.65 (0.24)	0.90 (0.15)*
	Pictures	Non rep	0.64 (0.17)	0.77 (0.16)*
		Rep	0.81 (0.18)	0.91 (0.11)*
False alarms	Words	Non rep	0.21 (0.19)	0.09 (0.12)*
		Rep	0.35 (0.23)	0.11 (0.17)*
		New	0.06 (0.08)	0.01 (0.02)*
	Pictures	Non rep	0.28 (0.23)	0.12 (0.09)*
		Rep	0.36 (0.23)	0.09 (0.13)*
		New	0.06 (0.11)	0.00 (0.01)*
Sensitivity (A′)	Words	Non rep	0.73 (0.13)	0.88 (0.11)*
		Rep	0.75 (0.13)	0.93 (0.11)*
	Pictures	Non rep	0.77 (0.12)	0.89 (0.07)*
		Rep	0.81 (0.13)	0.95 (0.04)*

Table 2. *Means (and SD) of hits, false alarms, and sensitivity index of the results of the associative recognition tasks broken down into remember and know judgments as a function of age groups, stimuli, and repetition conditions. Significant differences between age groups (p < 0.05) marked with **

			Older (n = 24)	Young (n = 28)
Remember responses				
Hits	Words	Non rep	0.31 (0.27)	0.56 (0.27)*
		Rep	0.48 (0.32)	0.78 (0.23)*
	Pictures	Non rep	0.40 (0.25)	0.59 (0.19)*
		Rep	0.60 (0.28)	0.82 (0.18)*
False alarms	Words	Non rep	0.12 (0.15)	0.03 (0.05)*
		Rep	0.22 (0.20)	0.02 (0.08)*
		New	0.03 (0.05)	0.00 (0.00)*
	Pictures	Non rep	0.16 (0.15)	0.04 (0.06)*
		Rep	0.21 (0.21)	0.03 (0.05)*
		New	0.03 (0.08)	0.00 (0.00)*
Sensitivity (A′)	Words	Non rep	0.71 (0.16)	0.86 (0.11)*
		Rep	0.74 (0.15)	0.93 (0.09)*
	Pictures	Non rep	0.74 (0.13)	0.88 (0.06)*
		Rep	0.79 (0.14)	0.94 (0.05)*
Know responses				
Hits	Words	Non rep	0.14 (0.16)	0.15 (0.14)
		Rep	0.18 (0.22)	0.12 (0.14)
	Pictures	Non rep	0.24 (0.21)	0.18 (0.18)
		Rep	0.21 (0.25)	0.10 (0.11)
False alarms	Words	Non rep	0.09 (0.13)	0.06 (0.09)
		Rep	0.13 (0.13)	0.09 (0.11)
		New	0.03 (0.04)	0.01 (0.02)
	Pictures	Non rep	0.13 (0.14)	0.08 (0.08)
		Rep	0.15 (0.16)	0.06 (0.10)*
		New	0.03 (0.05)	0.00 (0.00)*
Sensitivity (A′)	Words	Non rep	0.68 (0.14)	0.69 (0.13)
		Rep	0.69 (0.12)	0.63 (0.14)
	Pictures	Non rep	0.68 (0.13)	0.66 (0.14)
		Rep	0.65 (0.13)	0.65 (0.14)

conditions variable ($F_{2,100} = 52.36$, $p < 0.0001$, $\eta^2_p = 0.51$). Post hoc Bonferroni t-tests on this latter variable showed all the differences across the three means to be significant: 0.18 for unrepeated pairs, 0.23 for repeated pairs, and 0.03 for the new pairs. Moreover, the repetition conditions × groups interaction was also significant ($F_{2,100} = 13.54$, $p < 0.0001$, $\eta^2_p = 0.21$). Post hoc Bonferroni t-tests on this interaction showed that repetition increased the FA rates in older people (means of 0.25 and 0.36 for non-repeated and repeated pairs, respectively), but not in young people (means of 0.10 and 0.10 for non-repeated and repeated pairs, respectively), and older people made more FA on new pairs than young people did (means of 0.06, and 0.00, respectively). This significant interaction confirms the effect we wanted to elicit (e.g., Buchler *et al*., 2011; Light *et al*., 2004; Rhodes *et al*., 2008), and allows us to determine (by analyzing later the remember and know judgments) whether this effect is due to an incorrect use of familiarity, recollection, or both processes, the main goal of this study.

Finally, we calculated the memory index using the non-parametric sensitivity (A′) statistic (e.g., Snodgrass and Corwin, 1988). A mixed ANOVA 2 groups × 2 stimuli × 2 repetition conditions (see Table 1) on A′ showed that the main effect of the groups variable was significant ($F_{1,50} = 63.23$, $p < 0.0001$, $\eta^2_p = 0.56$), indicating that young people performed better than elderly people (means of 0.92 and 0.77, respectively), as was the main effect of the repetition conditions variable ($F_{1,50} = 8.54$, $p = 0.005$, $\eta^2_p = 0.15$), indicating that repeated stimuli were better recognized than unrepeated stimuli (means of 0.86 and 0.82, respectively), while the stimuli variable was marginally significant ($F_{1,50} = 3.85$, $p = 0.055$, $\eta^2_p = 0.07$, means of 0.86 and 0.82, for pictures and words, respectively). There were no significant interactions.

Remember responses

Second, we analyzed the remember (R) responses (see Table 2). With regard to hits, a mixed ANOVA of 2 groups × 2 stimuli × 2 repetition conditions showed that the main effects of the three variables were significant: groups ($F_{1,50} = 19.86$, $p < 0.0001$, $\eta^2_p = 0.28$), stimuli ($F_{1,50} = 4.64$, $p < 0.05$, $\eta^2_p = 0.09$), and the repetition conditions ($F_{1,50} = 88.21$, $p < 0.0001$, $\eta^2_p = 0.64$), indicating that young people made more R hits than elderly people (means of 0.69 and 0.45, respectively), pictures were better recognized by R judgments than words (means of 0.60 and 0.53, respectively), and repeated stimuli were better recognized by R responses than non-repeated stimuli (means of 0.67 and 0.47, respectively). No interactions were significant. These results coincide with those found in the literature, again supporting our procedure.

Regarding the FA on R responses (see Table 2), a mixed ANOVA of 2 groups × 2 stimuli × 3 repetition conditions (non-repeated, repeated and new) showed the main effect of the groups variable to be significant ($F_{1,50} = 45.09$, $p < 0.0001$, $\eta^2_p = 0.47$, indicating that elderly people made more FA on R judgments than young people: means of 0.13 and 0.02, respectively), as was the main effect of the repetition conditions ($F_{2,100} = 26.40$, $p < 0.0001$, $\eta^2_p = 0.35$). Post hoc Bonferroni t-tests on this latter variable showed the difference between the new pairs vs. non-repeated pairs to be significant (means of 0.02 and 0.09, respectively), as well as the difference between the new pairs vs.

repeated pairs (means of 0.02 and 0.12, respectively), while the difference between repeated and non-repeated pairs was not significant. The interaction groups × repetition conditions was also significant ($F_{2,100}$ = 14.61, p < 0.0001, η^2_p = 0.23). Post hoc Bonferroni t-tests on this interaction showed that repetition increases the R false alarm rates in elderly people (means of 0.14 and 0.22, for non-repeated and repeated pairs, respectively), but not in younger people (means of 0.03 and 0.03, for non-repeated and repeated pairs, respectively), and elderly people made more R false alarms on new pairs than young people did (means of 0.03, and 0.00, respectively). This significant interaction clearly confirms that the increase in FA produced by repetition in older people is due to the incorrect use of recollection, shedding light on the main hypothesis of our study. Below, in the analysis of FA on know judgments, we will analyze whether or not familiarity has a role to play in explaining this effect.

Finally, with regard to the A′ of the R responses (see Table 2), a mixed ANOVA of 2 groups × 2 stimuli × 2 repetition conditions showed the main effect of the groups variable to be significant ($F_{1,50}$ = 48.71, p < 0.0001, η^2_p = 0.49, indicating that young people performed better than elderly people on R responses: means of 0.91 and 0.74, respectively), and the repetition conditions variable also showed significant effect ($F_{1,50}$ = 22.27, p < 0.0001, η^2_p = 0.31, indicating that repeated stimuli were better recognized by R responses than non-repeated stimuli: means of 0.85 and 0.80, respectively). The stimuli variable was marginally significant ($F_{1,50}$ = 3.47, p < 0.07, η^2_p = 0.07, indicating a tendency toward a better performance on images than on words: means of 0.84 and 0.81, respectively).

Know responses

Third, we analyzed the know (K) responses (see Table 2). With regard to hits, a mixed ANOVA of 2 age groups × 2 stimuli × 2 repetition conditions showed the main effects and interactions of these three variables to be non-significant, which means that older and young people use familiarity in a similar way.

Regarding the FA on K judgments (see Table 2), a mixed ANOVA of 2 groups × 2 stimuli × 3 repetition conditions (non-repeated, repeated and new) showed that the main effect of the groups variable was significant ($F_{1,50}$ = 8.97, p = 0.004, η^2_p = 0.15, indicating that elderly people made more FA on K than young people: means of 0.09 and 0.05, respectively), and the repetition conditions variable was also significant ($F_{2,100}$ = 22.99, p = 0.0001, η^2_p = 0.32). Post hoc Bonferroni t-tests on this variable showed that the differences between the new pairs vs. non-repeated pairs were significant (means of 0.02 and 0.09, respectively), as were the differences between the new pairs vs. repeated pairs (means of 0.02 and 0.11, respectively), while the difference between repeated and non-repeated pairs was not significant. No interactions were significant, which, in relation to the main objective of our study, means that the increase in false alarms produced by the repetitions in elderly people is due to the incorrect use of recollection, and not familiarity, with implications that we discuss below.

Finally, with regard to the A′ of the K responses (see Table 2), a mixed ANOVA of 2 groups × 2 stimuli × 2 conditions showed that neither the main effects nor the interactions of these three variables were significant. However, comparisons of the A′ means of the young people (0.66) and elderly people (0.68) with the value 0.50 (which represents the performance at chance level) were both significant (t_{27} = 9.40, p < 0.0001, and t_{23} = 9.36, p < 0.0001, respectively), indicating that both populations use K responses efficiently and in a similar way to improve their performance, which seems to support the idea that familiarity does not decline with age.

DISCUSSION

As far as correct recognition is concerned, our results coincide with those obtained by other laboratories, confirming the appropriateness of our procedure. That is, young people recognize better than older people because the latter have more difficulties in encoding and remembering the contextual information associated with an item (confirming the associative-binding deficit hypothesis; e.g., Craik & Rose, 2012; Old & Naveh-Benjamin, 2008), while both samples use familiarity in a similar and effective way in recognition. On the other hand, the repetition of the pairs improves recognition in a similar way in both samples, which indicates that older people benefit from associative strengthening as much as young adults do. In other words, the repetition of a pair increases both the individual activation of the items involved and the strengthening of the association between them (e.g., Buchler *et al*., 2011), which proportionally improves its future retrieval in both young people and older people. However, the starting point for older people is lower, due to their aforementioned associative impairment. Finally, images tend to be recognized better than words in both samples (confirming the picture superiority effect; e.g., Defeyter *et al*., 2009), but this difference is significantly greater in older people than in young people, which experimentally supports using images as mnemonic aids in the elderly population and in people with cognitive impairment (e.g., Ally, Gold & Budson, 2009).

Regarding false recognition, our results show that the older people systematically made more FA than the young people, both on pairs of stimuli studied before but rearranged (e.g., Buchler *et al*., 2011; Light *et al*., 2004) and on new pairs of stimuli not studied before (illusions). Moreover, repetition does not increase false alarms in young people (confirming the idea that they correctly use a mechanism of *recall-to-reject*; e.g., Jones & Jacoby, 2001; Rotello & Heit, 2000), but it does increase them in older people. The traditional explanation for this difference is that older people use familiarity more than young people do (e.g., Buchler *et al*., 2011; Rhodes *et al*., 2008). However, and this is the novel result from our investigation, our results show that the origin of this increase lies in recollection, and not in familiarity. Applied to everyday life, our results would indicate that for both young and older people, seeing a person various times in the same context increases the future probability of correctly remembering the context in which they met this person. However, for older people (but not young people), seeing the same person several times in the same context also increases the probability of incorrectly remembering that they met this person in another similar but different context (which could have serious implications in eyewitness testimonies, for example). This *ironic* effect (Jacoby, 1999) seems to occur in situations that share similar perceptual and/or conceptual characteristics (Norman & Schacter, 1997; Schacter *et al*., 1997), as occurs in our experiment.

With regard to the result that older people regularly commit more FA on new stimuli (illusions) than young people, this has been considered by some authors (e.g., Abe, Fujii, Nishio *et al*., 2011; Hildebrandt *et al*., 2009a, 2009b) as an obvious symptom of cognitive impairment. Furthermore false recollections tend to increase proportionally to the degree of cognitive impairment (e.g., in patients with mild cognitive impairment, or Alzheimer disease).

Overall, these results cast doubts on the dual models of recognition, which are not able to explain how we can recollect something that we never studied (e.g., McCabe *et al*., 2009). In theory, we might think that our results fit the theoretical proposals of single-process models better, as they would predict that repetition would increase the strength of the items' trace, which would make them easier to recollect, provoking more correct recognitions as well as more incorrect ones (e.g., Tussing & Greene, 2001). However, at the same time, these models cannot explain why older people make more R false alarms than young people on new pairs of stimuli, as the trace strength of both stimuli is the same (in theory null) in both groups.

Our results seem to be more easily explained from the models related to the false memories' literature, supporting the idea that false remembering in older people seems be the result of both encoding difficulties and source misattributions at retrieval: what seems to occur in older people when they process stimuli that share perceptual and conceptual characteristics is that, first, they encode the stimuli through weaker episodic traces than young people, miscombining features from studied events (e.g., Treisman & Schmidt, 1982), due to their deficits in attention or executive control (associative-binding deficit hypothesis), and second, their retrieval can be affected by their own expectations or intuitions about what might have happened in the past, which can produce misrecollections or source monitoring errors. For example the activation/monitoring framework (e.g., Gallo, 2010, 2013; McDermott & Watson, 2001; Roediger, Watson, McDermott & Gallo, 2001) supports the idea that false recognition can be the result of either the automatic activation of lures at encoding (or at retrieval), or source monitoring errors committed at retrieval (guided by our expectations in memory decisions; see also Fandakova, Shing & Lindenberger, 2013; Johnson, 2006). Thereby, true and false memories would consist of reconstructions of what happened, arising false memories from the same source monitoring processes that produce true memories (which means that one can never be absolutely sure of the truth of any particular memory; Johnson, 2006). This theoretical framework could fit our results, in the sense that it could easily explain why older people committed more remember FA than young people, both in rearranged and in new pairs, and why older and young people do not differ in know FA, because they used activation in a similar way.

The activation/monitoring models explain the cause of these misrecollections in older people by an age-associated decline in the medial temporal lobe structures (hippocampus, perirhinal and parahippocampal cortices) and prefrontal cortex, which are very important for episodic binding and/or retrieval (e.g., Gallo, 2013; Mitchell & Johnson, 2009). Other authors (e.g., Hildebrandt *et al*., 2009a, 2009b; Romberg, McTighe, Heath *et al*., 2012) suggest that false memories could be due to the accumulation of plaques of the beta amyloid protein implicated in Alzheimer's disease (and in its prodromal condition, mild cognitive impairment) which leads to aberrant synaptic plasticity, thus increasing the likelihood of false recognition. As these plaques of the beta amyloid protein can be also found (by post-mortem autopsies) in the brains of some older healthy persons without dementia (e.g., Bennett, Schneider, Arvanitakis *et al*., 2006), this hypothesis could also explain why these people show false recollections. Thus, future research should combine behavioral experiments with neuroimaging and behavioral genetic measures to explain why people differ at encoding and recovering true and false information from their memory.

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. We acknowledge the collaboration of the undergraduate students Ana Burgos, Verónica Encinas, and Francisco Ródenas in selecting the experimental materials.

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.

Alameda, J. R. & Cuetos, F. (1995). *Diccionario de frecuencias de las unidades lingüísticas del castellano [Frequency dictionary of the linguistic units of Spanish]*. Oviedo: Servicio de Publicaciones de la Universidad de Oviedo.

Ally, B. A., Gold, C. A. & Budson, A. E. (2009). The picture superiority effect in patients with Alzheimer's disease and mild cognitive impairment. *Neuropsychologia*, *47*, 595–598.

Basile, B. M. & Hampton, R. R. (2013). Recognition errors suggest fast familiarity and slow recollection in rhesus monkeys. *Learning & Memory*, *20*, 431–437.

Bennett, D. A., Schneider, J. A., Arvanitakis, Z., Kelly, J. F., Aggarwal, N. T., Shah, R. C., *et al.* (2006). Neuropathology of older persons without cognitive impairment from two community-based studies. *Neurology*, *66*, 1837–1844.

Buchler, N. G., Faunce, P., Light, L. L., Gottfredson, N. & Reder, L. M. (2011). Effects of repetition on associative recognition in young and older adults: Item and associative strengthening. *Psychology and Aging*, *26*, 111–126.

Chen, T. & Naveh-Benjamin, M. (2012). Assessing the associative deficit of older adults in long-term and short-term/working memory. *Psychology and Aging*, *27*, 666–682.

Craik, F. I. M. & Rose, N. S. (2012). Memory encoding and aging: A neurocognitive perspective. *Neuroscience and Biobehavioral Reviews*, *36*, 1729–1739.

Defeyter, M. A., Russo, R. & McPartlin, P. M. (2009). The picture superiority effect in recognition memory: A developmental study using the response signal procedure. *Cognitive Development*, *24*, 265–273.

Dodson, C. S., Bawa, S. & Krueger, L. E. (2007). Ageing, metamemory, and high-confidence errors: A misrecollection account. *Psychology and Aging*, *22*, 122–133.

Fandakova, Y., Shing, Y. L. & Lindenberger, U. (2013). Differences in binding and monitoring mechanisms contribute to lifespan age differences in false memory. *Developmental Psychology*, *49*, 1822–1832.

Folstein, M. F., Folstein, S. E. & McHugh, P. R. (1975). "Mini-mental state": A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, *12*, 189–198.

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

Gallo, D. A. (2013). Retrieval expectations affect false recollection: Insights from a criterial recollection task. *Current Directions in Psychological Science*, *22*, 316–323.

Gallo, D. A. & Roediger, H. L. (2003). The effects of associations and aging on illusory recollection. *Memory & Cognition*, *31*, 1036–1044.

Hildebrandt, H., Haldenwanger, A. & Eling, P. (2009a). False recognition correlates with amyloid-$\beta_{1\text{-}42}$ but not with total tau in cerebrospinal fluid of patients with dementia and and mild cognitive impairment. *Journal of Alzheimer Disease*, *16*, 157–165.

Hildebrandt, H., Haldenwanger, A. & Eling, P. (2009b). False recognition helps to distinguish patients with Alzheimer's disease and amnesic MCI from patients with other kinds of dementia. *Dementia and Geriatric Cognitive Disorders*, *28*, 159–167.

Jacoby, L. L. (1999). Ironic effects of repetition: Measuring age-related differences in memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*, 3–22.

Johnson, M. K. (2006). Memory and reality. *American Psychologist*, *61*, 760–771.

Jones, T. C. & Jacoby, L. L. (2001). Feature and conjunction errors in recognition memory: Evidence for dual-process theory. *Journal of Memory and Language*, *45*, 82–102.

Kilb, A. & Naveh-Benjamin, M. (2011). The effects of pure pair repetition on younger and older adults' associative memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*, 706–719.

Koriat, A., Goldsmith, M. & Pansky, A. (2000). Toward a psychology of memory accuracy. *Annual Review of Psychology*, *51*, 481–537.

Light, L. L., Patterson, M. M., Chung, C. & Healy, M. R. (2004). Effects of repetition and response deadline on associative recognition in young and older adults. *Memory & Cognition*, *32*, 1182–1193.

McCabe, D. P., Geraci, L., Boman, J. K., Sensenig, A. E. & Rhodes, M. G. (2011). On the validity of remember–know judgments: Evidence from think aloud protocols. *Consciousness and Cognition*, *20*, 1625–1633.

McCabe, D. P., Roediger, H. L. III, McDaniel, M. A. & Balota, D. A. (2009). Aging reduces veridical remembering but increases false remembering: Neuropsychological test correlates of remember-know judgements. *Neuropsychologia*, *47*, 2164–2173.

McDermott, K. B. & Watson, J. M. (2001). The rise and fall of false recall: The impact of presentation duration. *Journal of Memory and Language*, *45*, 160–176.

Migo, E. M., Mayes, A. R. & Montaldi, D. (2012). Measuring recollection and familiarity: Improving the remember/know procedure. *Consciousness and Cognition*, *21*, 1435–1455.

Mitchell, K. J. & Johnson, M. K. (2009). Source monitoring 15 years later: What have we learned from fMRI about the neural mechanisms of source memory? *Psychological Bulletin*, *135*, 638–677.

Norman, K. A. & Schacter, D. L. (1997). False recognition in younger and older adults: Exploring the characteristics of illusory memories. *Memory & 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.

Paivio, A. (1971). *Imagery and verbal processes*. New York: Holt, Rinehart & Winston.

Rajaram, S. (1993). Remembering and knowing. Two means of access to the personal past. *Memory & Cognition*, *21*, 89–102.

Rhodes, M. G., Castel, A. D. & Jacoby, L. L. (2008). Associative recognition of face pairs by younger and older adults: The role of familiarity-based processing. *Psychology and Aging*, *23*, 239–249.

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 & Review*, *8*, 385–407.

Romberg, C., McTighe, S. M., Heath, C. J., Whitcomb, D. J., Cho, K., Bussey, T. J. & Saksida, L. M. (2012). False recognition in a mouse model of Alzheimer's disease: Rescue with sensory restriction and memantine. *Brain*, *135*, 2103–2114.

Rotello, C. M. & Heit, E. (2000). Associative recognition: A case of recall-to-reject processing. *Memory & Cognition*, *28*, 907–922.

Schacter, D. L., Koutstaal, W. & Norman, K. A. (1997). False memories and aging. *Trends in Cognitive Sciencies*, *1*, 229–236.

Snodgrass, J. & Corwin, J. (1988). Pragmatics of measuring recognition memory: Applications to dementia and amnesia. *Journal of Experimental Psychology: General*, *117*, 34–50.

Treisman, A. & Schmidt, H. (1982). Illusory conjunctions in the perception of objects. *Cognitive Psychology*, *14*, 107–141.

Tulving, E. (1985). Memory and consciousness. *Canadian Psychology*, *26*, 1–12.

Tussing, A. A. & Greene, R. L. (2001). Effects of familiarity level and repetition on recogntion accuracy. *The American Journal of Psychology*, *114*, 31–41.

Vannuci, M., Mazzoni, G., Marchetti, I. & Lavezzini, F. (2012). ''It's a hair-dryer. No, it's a drill''. Misidentification-related false recognitions in younger and older adults. *Archives of Gerontology and Geriatrics*, *54*, 310–316.

Wechsler, D. (2001). *Wechsler Intelligence Scale for Adults-III (WAIS-III)*. Madrid: TEA Ediciones.

Weinstein, Y. & Shanks, D. R. (2010). Rapid induction of false memory for pictures. *Memory*, *18*, 533–542.

Wixted, J. T. (2007). Dual-process theory and signal-detection theory of recognition memory. *Psychological Review*, *114*, 152–176.

Wixted, J. T. & Mickes, L. (2010). A continuous dual-process model of remember/know judgments. *Psychological Review*, *117*, 1025–1054.

Yonelinas, A. P. (2002). The nature of recollection and familiarity: A review of 30 years of research. *Journal of Memory and Language*, *46*, 441–517.

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.

Received 5 March 2014, accepted 4 September 2014