False recollection of emotional pictures in Alzheimer's disease

David A. Gallo a, b, *, Katherine T. Foster a, Jessica T. Wong a, David A. Bennett c, d

a University of Chicago, Department of Psychology, United States
b University of Chicago, Center for Cognitive and Social Neuroscience, United States
c Rush University Medical Center, Department of Neurological Sciences, United States
d Rush Alzheimer’s Disease Center, Chicago, IL, United States

A R T I C L E  I N F O

Article history:
Received 29 June 2010
Received in revised form 10 August 2010
Accepted 11 August 2010
Available online 19 August 2010

Keywords:
False memory
Emotion
Alzheimer's disease

A B S T R A C T

Alzheimer’s Disease (AD) can reduce the effects of emotional content on memory for studied pictures, but less is known about false memory. In healthy adults, emotionally arousing pictures can be more susceptible to false memory effects than neutral pictures, potentially because emotional pictures share conceptual similarities that cause memory confusions. We investigated these effects in AD patients and healthy controls. Participants studied pictures and their verbal labels, and then picture recollection was tested using verbal labels as retrieval cues. Some of the test labels had been associated with a picture at study, whereas other had not. On this picture recollection test, we found that both AD patients and controls incorrectly endorsed some of the test labels that had not been studied with pictures. These errors were associated with medium to high levels of confidence, indicating some degree of false recollection. Critically, these false recollection judgments were greater for emotional compared to neutral items, especially for positively valenced items, in both AD patients and controls. Dysfunction of the amygdala and hippocampus in early AD may impair recollection, but AD did not disrupt the effect of emotion on false recollection judgments.

© 2010 Elsevier Ltd. All rights reserved.

1. Introduction

Contrary to intuition, emotional events are not always remembered more accurately than neutral events. Emotionally arousing information can capture attention and activate specialized neural responses (e.g., amygdala modulation of hippocampus), and these processes can enhance memory for studied items (for reviews see Kensinger, 2009; Mather, 2007). However, emotionally arousing information also can be associated with elevated false recognition of nonstudied items and more liberal response bias (e.g., Dougal & Rotello, 2007; Herbener, Rosen, Khine, & Sweeney, 2007), potentially owing to greater conceptual relatedness between emotional items (e.g., Brainerd, Stein, Silveira, Rohenkohl, & Reyna, 2008; Maratos, Allan, & Rugg, 2000). Although different memory tasks tap these different effects of emotion, each pattern has been observed in both younger and older adults (e.g., Kapucu, Rotello, Ready, & Seidl, 2008; Kensinger, Brierley, Medford, Growdon, & Corkin, 2002), suggesting spared emotional processing with healthy cognitive aging.

Alzheimer’s disease (AD) tends to reduce the emotional memory effect for studied items, even though AD patients can have normal reactions to emotional stimuli (e.g., Hamann, Monarch, & Goldstein, 2000). Hamann et al. (2000) found greater recall for negative over neutral pictures in healthy controls, but not in AD patients. The AD group did show elevated recall of positive pictures over negative pictures in this study, but more recently, Kensinger et al. (2002) and Abrisqueta-Gomez, Bueno, Oliveira, and Bertolucci (2002) did not find a memory benefit for either positive or negative pictures in AD patients. Similarly, while some studies have shown emotional memory benefits for stories in AD patients (e.g., Kazui et al., 2000), a large-scale study by Kensinger, Anderson, Crowdon, and Corkin (2004) failed to find this effect. It is unclear why some studies have found emotional memory benefits in AD while others have not, but the effect tends to be reduced in AD. These reductions are often attributed to neuropathological changes in the amygdala (Kensinger et al., 2002; Mori et al., 1999), a common characteristic of early AD (Braak & Braak, 1991).

Surprisingly little is known about emotional false memory in AD, even though emotional memories are an important aspect of healthy aging (Carstensen et al., 2006). None of the aforementioned studies analyzed memory errors, and we know only three emotional memory studies in AD that have directly analyzed false recognition effects. LaBar et al. (2005) had subjects study neutral and negative emotional pictures from the International Affective Pictures Set (IAPS; Lang et al., 2005), and then
take an implicit memory test followed by an explicit recognition memory test (re-presenting pictures on the explicit test). They found greater false recognition of negative over neutral pictures on the explicit test in both AD patients and healthy older adult controls. 

Using this task, Gallo, Foster, and Johnson (2009) found that both younger and healthy older adults were more likely to incorrectly accept studied lures compared to nonstudied lures. Moreover, many of the incorrectly accepted studied lures were accompanied with relatively high levels of confidence. Because this task emphasized recollection-based responding, it was assumed that these high-confidence errors were at least partly based on false recollection. Critically, these errors were greater for positive and negative lures compared to neutral lures, suggesting elevated false recollection for emotional pictures in both age groups.

In the current study we used this task to test the effects of emotion on false recollection judgments in AD. Because AD patients tend to be more impaired in recollection than familiarity (e.g., Ally, Gold, & Budson, 2009), we simplified the task to help AD patients discriminate between studied targets and studied lures. The task was divided into three separate study/test blocks, corresponding to emotionally positive, negative, and neutral materials. We tested a group of patients in the early stages of AD as well as two control groups of cognitively healthy older adults. One control group was tested under identical conditions as the AD patients, in order to compare the effects of AD on task performance. Because the task was relatively easy for this control group, we also tested a separate control group under conditions of divided attention during study. This manipulation increased the overall level of memory errors in control subjects, thereby providing an additional opportunity to observe the effects of emotion on false recollection judgments.

3. Method

Participants. The participants included 18 AD patients from the Rush Alzheimer’s Disease Center (RADC, mean age = 77.3 years; SD = 7.5; mean education = 14.2 years, SD = 2.6; 11 females) and 18 healthy controls approximately matched on age, sex, and education (mean age = 72.8 years; SD = 7.8; mean education = 15.3 years, SD = 2.1, 12 females; all p’s > 0.05). We also tested 18 healthy controls in an unmatched divided attention condition (mean age = 79.2 years, SD = 7.7; mean education = 16.2 years, SD = 3.3, 11 females). These participants were similar in age to the AD patients, but had more years of education (p < 0.05). Patients were clinically diagnosed with probable AD based on NINCDS-ADRDA criteria (McKhann et al., 1984), following procedures reported by Bennett et al. (2006). Scores on the Mini Mental State Examination (Folstein, Folstein, & McHugh, 1975) were significantly lower in the AD patients (mean = 23.9, SD = 1.5, range = 22–27) compared to the matched controls (29.3, SD = 1.2, p = 0.001) and divided attention controls (28.8, SD = 1.0, p = 0.001), whereas the two control groups did not differ. (One control participant did not complete the MMSE.) Exclusion criteria included comorbid neurodegenerative disease, stroke, severe head trauma, cerebrovascular disease, alcohol or drug abuse, untreated depression, poor or uncorrected vision, or if English was not a primary language. Prior to the study phase, the AD patients and full attention controls rated their current mood as relatively positive (both means = 5.3 on a 1–7 scale). All participants (and caregivers, where appropriate) gave written informed consent and were paid, and review boards at the University of Chicago and the RADC approved the study.

Materials. Stimuli were 144 pictures from the IAPS. Based on norms (1–9 scale), 48 were neutral (range 4.18–5.99), 48 were negative (range 1.91–3.91), and 48 were positive (range 6.18–8.34). Negative and positive pictures were equally arousing (means = 5.62 and 5.58), and were more arousing than neutral pictures (3.31). Pictures of nudity and mutilation were excluded. Each picture was given a unique two to three word label (e.g., tourist with book for neutral images; toxic waste cleanup for negative images; astronaut in space for positive images). Although a variety of content was portrayed in each category (e.g., people, activities, objects), the emotional items were judged as more conceptually related than the neutral items in an independent rating task with older adults (see Gallo et al., 2009).

Procedure. All stimuli were presented via computer, and the experimenter entered responses. The task was divided into three study/test blocks, one block per emotional category (participants were counterbalanced). Within a block, the test phase immediately followed the study phase.

Each study phase contained 16 labels with their picture and 16 labels without their picture (randomly intermixed). In the full attention condition, item presentation was self-paced. On each trial, subjects were presented with a label, read it aloud, and rated it for emotional arousal (1 [low], 2 [medium], 3 [high]). They were then presented with the corresponding picture (if any) and rated it for arousal. In the divided attention condition, items were presented every second. Rather than making arousal judgments, subjects repeated random digits spoken every 2 s (not locked to stimulus onset). Errors on the digit task were rare (mean = 3.4 per block).
Each test phase used the labels as retrieval cues. Each test contained 16 labels that had been studied with their picture (targets), 16 labels that had been studied without their picture (studied lures), and 16 labels that were not studied (nonstudied lures). Items were counterbalanced across these conditions. Participants were instructed to respond “yes” to test labels only if they recall a corresponding picture, and “no” to labels that did not elicit such recollections, followed by a confidence judgment (1 [low], 2 [medium], 3 [high]). Test items were randomized and self-paced.

4. Results and discussion

Unless noted otherwise, results were significant at the conventional $p < 0.05$ (two-tailed), and effect sizes used $\eta^2_p$ (F-tests) or Cohen’s $d$ ($t$-tests).

**Study ratings.** Analysis of arousal ratings revealed an effect of item, $F(1, 34) = 13.31, MSE = 0.019, \eta^2_p = .28$, as pictures were more arousing than their labels (2.14 and 2.07), and an effect of emotion, $F(2, 68) = 49.14, MSE = 0.194, \eta^2_p = .59$, as negative items (2.40) were more arousing than positive (2.21), and each was more arousing than neutral (1.70, all $p$’s < 0.001). That older adults rated the negative images as more arousing than the positive images is inconsistent with the IAPS norms, which were based on younger adults, but this valence effect replicates the older adult ratings from Gallo et al. (2009) and was not associated with any systematic effects in false recollection judgments in that study. More importantly, there was no effect of group and no interactions in the current study, indicating that AD patients had normal reactions to the stimuli.

**Test performance.** Collapsing across the emotional conditions, the overall patterns of recollection judgments were very consistent (Table 1). Studied targets were endorsed more than studied lures in each group (all $p$’s < 0.001), demonstrating accurate recollection judgments even in AD patients, but studied lures were endorsed more than nonstudied lures in each group (all $p$’s < 0.05), demonstrating memory confusions. As expected, AD patients endorsed fewer targets and more lures (studied and nonstudied) compared to the other two groups, and similarly for divided compared to full attention in controls (all $p$’s < 0.05).

Table 1

<table>
<thead>
<tr>
<th></th>
<th>Older adults</th>
<th>AD patients</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Full attention</td>
<td>Divided attention</td>
</tr>
<tr>
<td>Neutral picture test</td>
<td>Neutral targets 0.88 (0.03)</td>
<td>0.74 (0.04)</td>
</tr>
<tr>
<td></td>
<td>Studied lures 0.07 (0.04)</td>
<td>0.24 (0.04)</td>
</tr>
<tr>
<td></td>
<td>Nonstudied lures 0.01 (0.00)</td>
<td>0.15 (0.03)</td>
</tr>
<tr>
<td>Positive picture test</td>
<td>Positive targets 0.85 (0.03)</td>
<td>0.82 (0.03)</td>
</tr>
<tr>
<td></td>
<td>Studied lures 0.12 (0.02)</td>
<td>0.37 (0.05)</td>
</tr>
<tr>
<td></td>
<td>Nonstudied lures 0.03 (0.01)</td>
<td>0.23 (0.04)</td>
</tr>
<tr>
<td>Negative picture test</td>
<td>Negative targets 0.89 (0.03)</td>
<td>0.69 (0.04)</td>
</tr>
<tr>
<td></td>
<td>Studied lures 0.05 (0.03)</td>
<td>0.28 (0.06)</td>
</tr>
<tr>
<td></td>
<td>Nonstudied lures 0.01 (0.00)</td>
<td>0.18 (0.05)</td>
</tr>
</tbody>
</table>

Note: standard errors are in parenthesis.

Analysis of studied targets revealed an effect of group and emotion, and an interaction between the two, $F(4, 102) = 2.66, MSE = 0.013, \eta^2_p = .09$. Positive targets were endorsed more than neutral or negative targets in divided attention controls, and more than negative targets in AD patients (all $p$’s < 0.05), with no other significant differences.

We also computed signal detection estimates of accuracy ($d’$), using the correction described in Snodgrass and Corwin (1988) to avoid ceiling and floor effects in the raw data. Our primary interest was in the ability to discriminate between studied targets and studied lures (Fig. 1). Each of these labels was presented during the study phase, but only the targets had been associated with a picture. We therefore assume that this discrimination was primarily driven by recollection of the picture for studied targets, although picture presentation also may have increased familiarity.

Consistent with prior work on recollection-based responding (Yonelinas, 2002), discrimination ($d’$) was considerably reduced by both divided attention and by AD in each emotional condition (all $p$’s < 0.001). Analysis of positive and neutral items confirmed an effect of emotion, $F(1, 51) = 11.57, MSE = 0.217, \eta^2_p = .19$, and group, $F(2, 51) = 48.65, MSE = 0.811, \eta^2_p = .66$, and no interaction. Overall discrimination was lower for positive (1.25) than neutral items (1.56). Similarly, discrimination was lower for positive (1.25) than negative items (1.47), $F(1, 51) = 6.62, MSE = 0.199, \eta^2_p = .12$, with an effect of group, $F(2, 51) = 45.55, MSE = 0.083, \eta^2_p = .64$. There also was a marginal interaction, $F(2, 51) = 2.83, MSE = 0.199, \eta^2_p = .10$, primarily driven by the lack of a difference between positive and negative items in the divided attention condition (see Fig. 1). Analysis of discrimination for neutral (1.56) and negative (1.47) items revealed only the effect of group, $F(2, 51) = 46.65, MSE = 0.953, \eta^2_p = .65$. Discrimination between studied targets and nonstudied lures also was computed, and the resulting analyses were mostly similar to those for studied lures. In addition to the effects of group (all $p$’s < 0.001), discrimination was lower for positive (1.70) than neutral items (1.89), $F(1, 51) = 4.37, MSE = 0.233, \eta^2_p = .08$, with no difference between positive and negative items (1.82) or between neutral and negative items, and no interactions.

Bias estimates ($C$) also were calculated using studied targets and studied lures, given that prior memory studies have found elevated bias estimates for emotional items (e.g., Dougall & Rotello, 2007). Analysis of neutral and positive items revealed only an effect of emotion, $F(1, 51) = 15.42, MSE = 0.099, \eta^2_p = .23$, as estimates were more liberal to positive (−0.19) than neutral items (0.04). Estimates also were more liberal to positive (−19) than negative items (0.08), $F(1, 51) = 24.83, MSE = 0.08, \eta^2_p = .33$, along with a group
interaction, $F(2, 51) = 4.75$, MSE = 0.08, $\eta^2_p = .16$. The interaction was driven by the lack of an effect in full attention controls, with significant effects in the other two groups (both $p < .01$). There were no differences in bias estimates for neutral (0.04) and negative (0.08) items. Bias estimates calculated with studied targets and nonstudied lures yielded similar results. Bias estimates were more liberal to positive (0.03) than neutral items (0.21), $F(1, 51) = 12.81$, MSE = 0.07, $\eta^2_p = .20$, with no other effects or interactions, and also were more liberal to positive (0.03) than negative items (0.25), $F(1, 51) = 17.71$, MSE = 0.07, $\eta^2_p = .26$, along with a group interaction, $F(2, 51) = 4.87$, MSE = 0.07, $\eta^2_p = .16$. The interaction again was driven by the lack of an effect in full attention controls, with significant effects in the other two groups (both $p < .01$). Bias estimates for neutral and negative items did not differ. As we elaborate in Section 5, the theoretical implications of these bias estimates are subject to interpretation, but in general these bias effects tracked the effects of emotion observed in false recollection judgments.

**Confidence judgments.** Confidence was analyzed for correct endorsements of targets and false endorsements of studied lures, excluding full attention controls and nonstudied lures owing to limited lure observations (Fig. 2). In general subjects were more confident for targets than lures, but this difference was reduced in AD patients. Moreover, both groups endorsed studied lures with medium to high levels of confidence (i.e., judgments greater than 2), suggesting some degree of false recollection. Analysis of targets revealed only an effect of group, $F(1, 34) = 8.72$, MSE = 0.458, $\eta^2_p = .204$, indicating that AD patients were less confident in their recollection judgments compared to controls. Analysis of studied lures revealed a trend for an effect of emotion, $F(2, 52) = 2.44$, MSE = 0.120, $\eta^2_p = .09$, $p = 0.097$, which primarily was driven by significantly greater confidence for negative over neutral items, $t(27) = 2.24$, SEM = 0.091, $d = 0.37$.

**5. Discussion**

AD patients showed a similar emotional effect on false recollection judgments as healthy controls, despite considerably reduced memory for studied pictures. Emotional effects on memory for studied information can be reduced in AD patients (e.g., Kensinger et al., 2002), potentially due to dysfunction of the amygdala and hippocampus, but the early stages of AD do not appear to disrupt the effect of emotional content on false recollection judgments. These effects may have significant clinical implications because many important life events recalled by older adults are emotional (e.g., Berntsen & Rubin, 2002). To the extent that AD spares emotional processing of these events, but limits the ability to monitor memory for accuracy, AD patients may be especially prone to emotional autobiographical memory distortion (see Budson et al., 2004).

The current results are consistent with prior AD studies that reported elevated false recognition of emotional compared to neutral items (Brueckner & Moritz, 2009; Labar et al., 2005), and they extend these findings to false recollection judgments. Our results also are consistent with the valence effects of Brueckner and Moritz (2009), who found greater effects of positive over negative words in AD patients. (LaBar et al. only used negative items.) We found valence effects in false recollection judgments for AD patients and healthy controls (positive > negative), and we also found that positive targets were more likely to be correctly recognized than negative targets in divided attention controls and in AD patients. In contrast, Budson et al. (2006) found no emotional effect on false recognition in AD patients, although there was a trend for more liberal bias estimates in controls. Numerous methodological differences may underlie this discrepancy, but it is worth reiterating that we used highly emotional pictures, and even with these stimuli the false memory effects were stronger for positive than for negative items. Thus, the use of relatively less arousing word lists and only negative items may have limited the ability of Budson et al. to detect emotional memory effects in AD patients.

Our finding that positive valence influenced both true and false responses is consistent with the source monitoring framework, which assumes that the same biases and confusions that drive memory errors also can influence responding for studied items (Johnson, 2006). These effects also may be related to a more general preference for positive information with aging (Carstensen, Mikels, & Mather, 2006). However, such positivity effects are not always found (see Murphy & Isacowitz, 2008), and in the current study, negative information enhanced confidence judgments to errors for studied lures. Considered overall, our findings suggest that both positive and negative pictures can enhance false memory effects (also see Gallo et al., 2009).

Emotion might enhance false recollection judgments by increasing conceptual relatedness between items. Enhanced relatedness could result from item-specific semantic associations or from the emotional category itself (e.g., happy items, sad items, etc.). Our results are more consistent with the latter, because AD patients showed similar effects of emotion on memory errors as healthy controls. If semantic associations were the critical factor then these effects should have been reduced in AD, analogous to AD effects on association-based false recognition (e.g., Budson et al., 2006; Gallo et al., 2006). In contrast, both groups were sensitive to the emotional category of the items, as indicated by their similar arousal responses at study. Further support for an emotional category effect comes from studies demonstrating liberal bias estimates and enhanced false recognition for emotional items even after equating for item associations or content (e.g., Herbener et al., 2007; Kapucu et al., 2008).

In addition to false recollection judgments, we also found effects of emotion on estimates of response bias. These bias estimates tended to track the emotional effects observed on false recollection judgments, with more liberal bias estimates for emotional items. These effects are consistent with prior results (e.g., Dougal & Rotello, 2007; Gallo et al., 2009; Kapucu et al., 2008), and they extend these findings to AD patients. However, these effects should be interpreted with caution for two reasons. First, because our primary interest was in false recollection judgments, we tested more lures than targets. This factor was held constant across our emotional conditions and therefore is unlikely to have caused the observed effects of emotion. Nevertheless, including more lures may have had an overall effect on bias estimates. Second, bias estimates are based on the hypothetical relationship between a response criterion and the memory distributions for targets and lures. As a result, changes in bias estimates might be caused by the creation of false memories, which could influence the lure distribution instead of the response criterion (see Wixted & Stretch, 2000).
emotional content increases false recollections for test lures, then this could increase false recollection judgments and corresponding estimates of bias independent of changes in response criterion. Thus, the observed effects of emotion on bias estimates are theoretically ambiguous. They may be driven by differences in response criterion or false recollection.

Our task emphasized recollection by using verbal retrieval cues and explicit picture recollection instructions, and the finding that false recollection judgments were made with relatively high confidence under these conditions suggests that these memory errors were at least partly based on the subjective experience of false recollection. Evidence that our task involved recollection-based responding was that accuracy was considerably reduced by both divided attention at study and by AD, both of which tend to affect recollection more than familiarity (Yonelinas, 2002). This is not to say, though, that familiarity plays no role in the subjective experience of false recollection or in false recollection judgments. According to the familiarity/corrobotation hypothesis (Lampinen, Neuschatz, & Payne, 1999), the familiarity of misleading retrieval cues may motivate subjects to search memory for corroborating evidence, causing them to reconstruct a false recollection from partially recollected features. Emotion may enhance this reconstructive process by increasing conceptual similarities between studied and nonstudied items, thereby enhancing familiarity as well as the availability of potentially confusable features in memory. AD also may enhance this reconstructive process, by degrading recollection and increasing reliance on familiarity, although these effects may be offset by reduced recollection of the features contributing to reconstruction. To more fully understand the processes that contribute to false recollection, and how Alzheimer’s disease may affect these processes, additional research is needed that uses tasks emphasizing recollection-based responding.

Acknowledgements

We are grateful to Celia Berdes, Barbara Eubeler, Raj Shah, and Presbyterian Homes. This work was supported by National Institute of Aging grants AG13045 and P30AG10161.

References


