RUNNING HEAD: PRIMING PROBLEM SOLUTIONS WITH FALSE MEMORIES

Priming Older Adults and People with Mild to Moderate Alzheimer's Disease Problem-solving with False Memories

Mark L. Howe[1] and Shazia Akhtar[2]

[1]Department of Psychology, City, University of London

[2]School of Life and Medical Sciences, University of Hertfordshire

Authors' Note

Mark L. Howe: http://orcid.org/0000-0002-5747-5571

Shazia Akhtar: http://orcid.org/0000-0002-1064-7743

The authors have no conflicts of interest to disclose.

Correspondence concerning this research should be addressed to Mark L. Howe, Department of Psychology, City, University of London, EC1V 0HB, United Kingdom; e-mail: Mark.Howe.1@city.ac.uk.

IN PRESS: *Cortex*

Author Contribution Statement

Mark L. Howe: Conceptualization; Formal analysis; Investigation; Methodology; Project administration; Resources; Supervision; Validation; Roles/Writing – original draft; Writing – review & editing.

Shazia Akhtar: Conceptualization; Formal analysis; Investigation; Methodology; Project administration; Resources; Supervision; Validation; Roles/Writing – original draft; Writing – review & editing.

Abstract

In two experiments we investigated whether older adult controls (OACs) and people with mild and moderate Alzheimer's disease (AD) benefit from false memory priming effects in subsequent problem-solving tasks. In addition, and unlike in previous false memory priming studies with older adults, we examined latency measures in the recognition phase. In Experiment 1 participants were asked to solve compound remote associate task (CRAT) problems, half of which had been preceded by the presentation of Deese/Roediger-McDermott (DRM) lists whose critical lures (CLs) were also the solutions to those problems. In Experiment 2, we used a similar paradigm but investigated whether CLs could prime solutions to subsequent analogical reasoning problems. In this latter experiment, we also examined whether these priming effects were stronger when the activation of the CL term occurred during the memory task (was presented as part of the list; i.e., true memories) or when these items were not presented but arose during encoding due to spreading activation (i.e., false memories). We found that all three groups' performance on these tasks was facilitated only by false memories spontaneously generated from the prior presentation of DRM lists. That is, performance on CRATs and analogical reasoning tasks was better (greater accuracy and faster speed) when those problems were preceded by DRM lists whose CLs also served as the solution to those problems. These findings are consistent with previous results from studies with children, young adults, and older adults and extends them to people with more moderate AD.

Keywords: Alzheimer's disease, analogical reasoning, compound remote associates, DRM paradigm, false memory priming.

Introduction

Alzheimer's disease (AD) is a progressive neurodegenerative condition which is characterized by a reduction in learning and memory performance as well as rapid forgetting of new information. Long-term memory is the most profoundly affected cognitive domain in AD. In particular, research suggests (e.g., Moulin, Perfect, & Jones 2000) that an encoding deficit, rather than increased forgetting, underlies the poor performance on episodic memory tasks for people with AD.

It has been suggested that one reason for episodic memory deficits in aging generally (e.g., Dew & Giovanello, 2010; Naveh-Benjamin, Shing, Kilb, Werkle-Bergner, & Lindenberger, 2009), and Alzheimer's disease more specifically (e.g., Fleishman et al., 2005; Meiran & Jelicic, 1995), is an impairment in strategic recollection. That is, age differences will be maximized when memory tasks rely on consciously-controlled processes. In other words, tasks that require explicit memory are at greater risk of exhibiting age- and AD-related declines than tasks that require less-effortful, implicit memory. Whereas explicit memory tasks require participants to consciously recall or recognize recently processed information, implicit memory tasks do not. Instead, implicit memory is indexed by finding that prior exposure to information can improve task performance even in the absence of participants' explicit awareness or memory for the information they were exposed to earlier.

One way of demonstrating this episodic deficit is with tests of recognition, where participants have to differentiate old items (what they have previously studied) from new items (what they did not study earlier). In this case, participants with AD tend to correctly recognize fewer old items than controls. They also make more false positive errors (incorrectly judging a new item to be old) than healthy older control participants. This combined problem of relatively low hits and high false positives is interesting because it suggests that as well as merely forgetting previously presented stimuli, AD patients also confuse new items as old.

The Deese Roediger-McDermott (DRM: Deese, 1959; Roediger & McDermott, 1995) paradigm is commonly used to examine false memories. Participants are presented with a list of words all of which are associatively related (e.g., *table*, *sit*, *legs*) and are also associated with a critical lure (CL) word that is never presented (i.e., *chair*). When memory is subsequently tested using recall or recognition tasks, healthy participants show a tendency to falsely recall and recognize the CLs as having been presented on the list (e.g., Balota, Cortese, Duchek, Adams, Roediger, McDermott, & Yerys, 1999; McDermott, 1999).

Spreading associative activation theories (e.g., Associative-activation theory, AAT, Howe, Wimmer, Gagnon, & Plumpton, 2009; Activation-monitoring theory, AMT, Roediger, Watson, McDermott, & Gallo, 2001) are the more dominate explanations of the increase in false positives for CLs. These theories suggest that false memories are produced because of two distinct processes: (1) an activation process where items in associative memory become active upon presentation and this activation then spreads to related, but non-presented, information in memory and (2) for adults at least, there is an additional source-monitoring process (see AMT in particular) that occurs during retrieval. Because the presentation of each of the list words during the DRM task automatically activates the related but unpresented CL, the CL is activated many times through an automatic spread of activation within the associative network. It is this activation that increases the feeling of familiarity for the item, while at the same time reducing the ability to remember the source of its activation (source-monitoring process). This monitoring process involves a two-step model of recognition (e.g., Atkinson & Juola, 1974), where there are separate criteria for *old* and *new* responses. If an item's familiarity falls above the higher criterion, subjects can make a rapid *old* response. If familiarity falls below the lower criterion, subjects will make a rapid *new* response. Items whose familiarity falls somewhere in between the two criteria undergo an additional checking process, yielding longer latencies. The heightened familiarity of the CL should lead participants toward endorsing the CL as *old*, but to do this, they will likely engage the slower checking process. This is because although the CL appears familiar with multiple activations, it was never actually presented. Thus, responses to CLs should be slower than responses to unrelated or weakly related items, items that should be less likely to engage the additional checking process.

The majority of studies using the DRM paradigm have reported accuracy not speed data, with latencies having only rarely been reported. One of these exceptions appears in a study by Jou, Matus, Aldridge, Rogers, and Zimmerman (2004, Experiment 1) who showed reaction times for false alarms were reliably longer than reaction times for hits. However, Tun, Wingfield, Rosen, and Blanchard, (1998) reported no difference between hits and false alarm RTs. Coane, McBride, Raulerson, and Jordan (2007) in two experiments using short-term memory tasks reported longer RTs for false alarms compared to hits. Interestingly, in their study they reported slower RTs for correct rejections compared to weakly related items.

Importantly, false memories generated in the DRM paradigm do not simply represent false positive errors on a memory test but can also be considered a kind of self-generated memory response – that is, it is information generated from the activation of one’s own internal semantic network. As such, this self-generated information can be more powerful and more memorable than information that was actually presented (other-generated information or “true” memories). Indeed, it is well known that self-generated information is less susceptible to the effects of forgetting, being better retained over long retention intervals (e.g., see Mulligan & Lozito, 2004; Sui & Humphreys, 2015).

This interpretation concerning the importance of self-generation effects in memory (and its extension to false memories) has gained considerable currency in the child development literature (e.g., Cunningham, Brebner, Quinn, & Turk, 2014; Ford & Lobao, 2018), the literature on memory in young adults (e.g., Mulligan & Lozito, 2004) as well as old adults (for a review, see Gutchess & Kensinger, 2018), and quite possibly could extend to older adults with AD (at least in the early stages) (Klein, Cosmides, & Costabile, 2003; Rosa, Deason, Budson, & Gutchess, 2015). Because the self can be the very "glue" that binds encoded elements together to create strong and durable traces, these traces are particularly well remembered (e.g., Humphreys & Sui, 2016; Sui & Humphreys, 2015).

Moreover, it turns out that self-generated information serves as better primes in subsequent problem-solving tasks then other-generated information, especially following a delay (see Howe, Wilkinson, & Monaghan, 2012). For example, Howe, Threadgold, Norbury, Garner, and Ball (2013) presented children and adults with DRM lists that either contained the CLs as presented items (true memory) or used the standard DRM procedure where the CLs were not contained in the list (false memory). The CL was also the solution to a subsequent series of analogical reasoning problems. Verbal proportional analogies refer to problems of the sort, *a* is to *b* as *c* is to (*d*)? where the participant's job is to come up with the d term. As a specific example, participants might be presented with a problem such as "Wardrobe is to Clothes as Bed is to ?" In this case, the question is whether the solution word "sleep" can be primed by the DRM list whose CL is "sleep."

A further question was whether the CL, regardless of whether it was actually presented or self-generated, could serve to prime the answers to these reasoning problems. The results showed that only the self-generated (false memory) CLs from DRM lists served as effective primes for these reasoning problems. That is, both children and adults were more likely to solve these reasoning problems, and solved them more quickly, when the CL had been self-generated and that the actually presented CLs had no beneficial effect on problem-solving. Indeed, for both age groups, there was no significant difference between solution times for unprimed analogical problems and those primed by true memories.

Recently, we (Akhtar & Howe, 2019) replicated this same finding with healthy older adults and those with mild AD. Although as expected, healthy older adults were better (both in terms of speed and accuracy) at solving problems than people with mild AD and both groups performed better when false memories served as primes than when true memories served as primes or when there were no primes. Like the data from both children and adults, there were no reliable differences between unprimed and true primed problems for either the healthy older adults or those with mild AD.

False memories not only serve as effective primes for analogical reasoning problems but also for compound remote associates tests (CRATs). CRATs are problems involving the presentation of three unrelated words (e.g., *walk, beauty, over*) and the participant has to come up with a fourth word (in this case the CL, *sleep*) that when combined with each of the presented words, creates a new compound word or phrase (e.g., *sleep walk, beauty sleep, sleep over*). Howe, Wilkinson, Garner, and Ball (2016) examined these problems with children (also see Howe, Garner, Charlesworth, & Knott, 2011) and adults. Like the findings for analogical reasoning problems, they found that DRM lists whose CLs were the solution to a series of subsequently presented CRATs were solved more often and more rapidly when the solution CL had been falsely generated. Similarly, Akhtar, Howe, and Hoepstine (in press) found that healthy older adults and those with mild AD solved CRATs more frequently and more rapidly when DRM lists whose CL had been falsely recognized prior to the presentation of the CRATs than when they had not been falsely recognized or when they had not been preceded by a relevant DRM list.

In the current research, we wanted to (1) replicate our previous findings on false memory priming effects in healthy older adults and those with mild AD, (2) extend these findings to participants with more moderate AD, and (3) explore whether latencies in the recognition phase can differentiate responses to actually presented list items, false memory items generated by associative activation during encoding (related but unpresented items), and foil items that were never presented and were unrelated to list items. We did this because, first, we wanted to see the extent to which self-generated information facilitates problem-solving performance in people with moderate AD not just mild AD. Second, there exists a large clinical literature showing that there are a number of important semantic deficits associated with AD, particularly of the moderate than mild variety (e.g., Hodges, Salmon, & Butters, 1992; Howard & Patterson, 1992). However, despite these deficits, we have found that priming effects exist at least in older people with mild AD (Akhtar & Howe, 2019; Akhtar et al., in press) and Nebes (1989) as well as Evard, Colombel, Gilet, and Corson (2016) have shown that semantic priming effects in AD remains relatively intact. What we are addressing in the following two experiments, is whether priming effects are still present in more severe cases of AD. That is, our current experiments are designed to examine whether semantic associative processes are still engaged automatically in participants with moderate AD and whether the results of these processes can be used to facilitate subsequent problem solving.

We made the following predictions. First, although healthy older adults and adults with mild AD can make appropriate responses on recognition tests (e.g., consciously discriminate between actually presented items and unrelated foils), we anticipate that those adults with moderate AD will not be able to make these conscious recognition responses, instead claiming that all recognition items are “new” (Akhtar, Moulin, & Bowie 2005; Moulin 2002; Moulin et al., 2000). Second, despite their inability to make appropriate verbal responses on the recognition test, we anticipated that the time it takes adults with moderate AD to make a response will follow the same trajectory as that for healthy older adults and those with mild AD. That is, as discussed earlier, when reaction times have been reported for participants in the DRM task, the lag for responding to CLs tends to be the longest followed by quicker response times to items actually presented (e.g., Coane et al., 2007; Jou et al., 2004). Thus, regardless of the participant's actual verbal or conscious recognition response, we predict that the pattern of response times will be similar across all three groups of participants. Of course, there will be group differences in the absolute values of the response times, where people with moderate AD will take longer overall to respond than people with mild AD who in turn will take longer to respond than healthy older adults.

Finally, if these patterns hold across groups, then it may be the case that participants in all groups will have formed false memories for the CL and these false memories should facilitate performance on subsequent problem-solving tasks (CRATs in Experiment 1 and analogies in Experiment 2). Thus, although people with moderate AD may not "know" they have formed a false memory for the CL, the response latencies may show that one has formed automatically in memory, indicating that some associative networks are still intact. If so, then there should be a performance advantage for those problems whose solutions have been primed by that CL. Note that no part of these studies or analyses were pre-registered prior to the research being conducted. We report how we determined our sample size, all data exclusions (if any), all inclusion/exclusion criteria, whether inclusion/exclusion criteria were established prior to data analysis, all manipulations, and all measures in the study. The anonymized data and materials for both experiments have been publicly archived at the University of Hertfordshire Research Archive and can be accessed via https://doi.org/10.18745/ds.22095.

Experiment 1

In our first experiment, we wanted to establish whether older adult control (OAC) participants, people with mild AD, and people with moderate AD could generate CLs from

DRM lists and subsequently use them to facilitate problem-solving. To do this, we examined people's ability to remember DRM lists and self-generated CLs from those lists and their ability to use those false memories to aid their performance on subsequent CRAT problems.

Method

Participants. Using G*Power (Faul, Erdfelder, Lang, & Buchner, 2007), we ran an a priori power analysis where we estimated the effect as medium (d = 0.5) and power of 0.80. The power analysis showed that we needed a sample of 30 participants per group. A sample of 90 participants was recruited whose demographic and other characteristics are shown in Table 1. This study was approved by the psychology ethics committee at City, University of London (ethics approval reference PSYETH (T/F) 16/17 209).

Sixty participants had a clinical diagnosis of probable or possible AD (McKhann, Drachman, Folstein, Katzman, Price, & Stadlan, 1984). Of these 60 participants with AD, 30 had mild AD and 30 had moderate AD. Participation took place in the participant's home or at a local Day Centre to the participant. Thirty participants made up the OAC group. These people were community dwelling and were recruited from a panel of older adults who had expressed an interest in participating in research. Participation took place at City, University of London. There were no significant differences between OACs and AD participants in terms of age, years of education, or performance on the National Adult Reading Test, but as shown in Table 1, all three groups differed on most cognitive tests.

INSERT TABLE 1 ABOUT HERE

Design, Materials, and Procedure. A 3 (Group: Moderate AD vs. Mild AD vs. OAC) x 3 (Priming: prime/FM vs. prime/noFM vs. unprimed) mixed design was used, where the first factor was between-participants and the second was a within-participant factor. We followed the same procedure as Akhtar et al. (in press) and Howe et al. (2011, 2016) such that each participant was primed on half the subsequent CRAT problems with preceding DRM lists whose critical lures were also the solution to those CRAT problems. The order of both the DRM lists and CRAT problems was carefully counterbalanced to eliminate order effects. Ten CRAT problems were selected from the normative data collected previously by Akhtar et al. (in press) (see Appendix A). Their unprimed solution rates ranged between 30% and 80% in order to ensure that they were neither too easy nor too hard. In addition, 10 DRM lists whose CLs were the corresponding solutions to these CRATs were selected. Each list contained 10 word associates of the critical lure and were taken from the Roediger, Watson, McDermott, and Gallo (2001) norms. Associative words that overlapped with the items in the CRAT problems were removed and replaced with another associate. Lists were randomly divided into two groups of five with participants being primed using one of these sets. In order to prevent differences in false memory rates, the two sets of five DRM lists were equated on backward associative strength (BAS) (List Set 1 BAS = 0.777, List Set 2 BAS = 0.725).

Participants were given one set of the five DRM lists in a randomized order. Each list was presented verbally, followed by a distractor task (counting backwards in 3s). This was followed by a recognition test whereby participants were verbally presented, in random order, the 5 critical lure words from the studied DRM lists, 5 unstudied and unrelated critical lures, 32 true items from the studied DRM lists, and 32 foils unrelated to studied DRM lists. A recognition test was implemented rather than a recall test to reduce effects of priming during retrieval (Olszewska & Ulatowska, 2013). For each word presented in the recognition task, participants had to select on the keyboard either [O], indicating that the word was Old and that they recognize the word from the previously presented lists, or [N] if they thought the current word presented was a new word that they did not hear in the previous word lists. Following completion of the DRM lists, participants had to solve CRAT problems. Participants were told that they would see three words and that they should try and produce a fourth word that, when combined with each of the three items, would make up a common compound word or phrase. Participants were first given three demonstrations by the experimenter followed by two practice problems prior to the experimental CRATs. The three problem words were presented on a computer screen simultaneously in a horizontal orientation, with one word above, below, and at the center fixation point. Participants were given 30s to produce the solution (this was a verbal solution and participants first response was recorded). If the solution was produced within the time limit, both the solution word and solution time were recorded, and the next problem was presented. If participants did not produce the correct response within the time limit, the solution was provided by the experimenter and the program automatically moved to the next problem. All 10 CRAT problems were completed, and solution rates and solution times were collected for correctly solved CRAT problems.

Results and Discussion

Recognition Task

We were interested in two factors in the recognition task. First, we were interested in recognition rates and, more specifically, whether all three groups produced false memories. In addition, we wanted to compare recognition rates across the different item types. Second, we were interested in latency measures for correctly recognized items across different item types.

Recognition Rates. As predicted the moderate AD group responded ‘N’ (new) to all items, indicating they were not able to consciously discriminate between the different item types. Because of this, we removed them from this set of analyses. The mean correct recognition rate (percentage) and the mean recognition reaction time (milliseconds) were calculated for each participant and analyzed separately in a series of 2(Group: Mild AD vs. OAC) x 4(Item Type: Critical lures vs. Unrelated Critical Lures vs. List Items vs. Foils) analyses of variance (ANOVAs). The recognition task (see Table 2) showed that both the OACs and the Mild AD group created false memories for the critical lure words, with the mild AD group falsely recognizing the critical lure 64.93% ($M = 3.3$, $SD = 1.31$) of the time and the OAC group 63.01% ($M = 3.2$, $SD = .91$) of the time. There was no reliable difference.

The 2 x 4 ANOVA revealed a significant main effect of item type, F (3, 58) = 1175.24, $p < .001$, $\eta^2_p = .953$. Pairwise comparisons revealed greater recognition for list items (79%) compared to correct rejection of foil items (61%) ($M = 23.40$, $SD = 3.54$ vs. $M = 19.53$, $SD = 4.78$, respectively) and greater false recognition of CL words (64%) compared to unrelated CL words (40.4%) ($M = 3.03$, $SD = .80$ vs. $M = 2.02$, $SD = 1.08$, respectively). There was also a main effect for group F (1, 58) = 63.7, $p < .001$, $\eta^2_p = .53$. To investigate this further, we conducted separate t-tests for item types. For actual list items, the OAC group correctly recognized significantly more items ($M = 27.28$, $SD = 3.16$) than people with Mild AD ($M = 23.1$, $SD = 3.15$), t (59) = 3.5, $p < .001$. Concerning foil items, again the OAC group correctly rejected significantly more items ($M = 22.43$, $SD = 4.54$) than people with mild AD ($M = 16.63$, $SD = 2.94$), t (59) = 2.28, $p < .05$. There was no interaction ($F < 1$).

INSERT TABLE 2 ABOUT HERE

Signal Detection Analysis. The values of *d'* and *C* are shown in Table 3. Again, the moderate AD group were excluded from this analysis as they responded ‘N’ to all items in the recognition task. Signal detection measures for hits and critical lures were analyzed in separate 2 (Group: Mild AD vs. OAC) X 2 (Item Type: critical lures vs. studied items) mixed-model ANOVAs. The analysis of *d'* revealed a main effect of item type, where discriminability was better for critical lures compared to list items, F (1, 58) = 68.72, $p < .001$, $\eta^2_p = .23$. There was no main effect of group and no interaction (*p*s ns).

Analysis of the criterion C revealed a more liberal bias for the critical lures F (1, 58) = 10.15, $p < .001$, $\eta^2_p = .48$ compared to studied items. There was also a main effect of group F (1, 58) = 23.59, $p < .001$, $\eta^2_p = .41$ where people with Mild AD revealed a more liberal bias compared to OACs. There was no significant interaction.

INSERT TABLE 3 ABOUT HERE

Recognition Times. Despite the moderate AD group not being able to consciously discriminate between the item types, we analyzed their latencies during the recognition phase for the different item types. We conducted these analyses in order to create a proxy measure of recognition memory for these participants. That is, although people with moderate AD may not be able to provide explicit recognition responses, perhaps any differences in their latencies to say "new" to the various items on the test list reflects their implicit recognition response. For example, perhaps longer latencies to say "new" to "old" items reflect the fact that there is a residual trace for that item in memory, one that inhibits or delays the incorrect overt "new" response. If so, then perhaps shorter latencies reflect the absence of such a trace for truly "new" items. Finally, if false memory probes exhibit a response delay similar to that seen for items that had actually been presented, then this may indicate that traces also exist in memory for critical lures. What this would show is that even for people with moderate AD, spreading activation mechanisms remain intact for semantic associations in long-term memory.

To do this we first looked at the pattern of correctly recognized item RTs across the different item types for the OACs and the mild AD group and then compared their patterns to the moderate AD group. Data points with RTs that fell above or below 2 *SDs* for each participant were omitted from the analysis. The mean recognition reaction times (milliseconds) were calculated for each participant and analyzed in a 2 (Group: OAC vs Mild AD) vs 4 (Item type: CLs, List Items, Unrelated unpresented CLs and foils) mixed-model ANOVA (see Table 4). There was a main effect of item type [F (3, 57) = 194.2 $p < .001$, $\eta^2_p = .77$]. Tukey's post-hoc tests indicated longer RTs for CLs (6373 ms, $p < .05$), followed by list items (5298 ms, $p < .05$), foils (3737 ms), and unrelated CLs (3876 ms) (there were no reliable difference between the latter two). As expected there was a main effect of group F (1, 58) = 427.9, $p < .001$, $\eta^2_p = .88$, Tukey's post-hoc tests showed the mild AD group were reliably slower (5428.2 ms) compared to the OACs (4214.6 ms). There was no interaction ($p > .05$). We ran separate analyses comparing false recognition of CLs versus correct rejections of CLs [F (1, 58) = 347.59 $p < .001$, $\eta^2_p = .77$] and found that false recognition of critical lures was reliably slower (6373 ms) than correct rejections of these same lures (5149 ms). We also compared correct recognition of list items to those list items that were not correctly recognized F (1, 58) = 369.78, $p < .001$, $\eta^2_p = .86$, and found that correctly recognized list items were reliably slower (5298 ms) than those list items that were not correct recognized (3938 ms). There was no difference in reaction times for correctly rejected foils compared to foils not rejected ($p > .05$) and no difference in reaction time between correctly rejected unrelated CLs and those that were not rejected.

Next, we examined whether this pattern of latencies was also evident in the moderate AD group. If so, then based on their item type latencies, we could work out how many *probable* CLs, correct rejections, and list items moderate AD people *implicitly* recognized (see Table 2). Because there were no differences in RT for correctly rejected foils/unrelated

CLs versus nonrejected foils/unrelated CLs for the Mild AD group and the OACs, we only included the reaction times for foils and unrelated CLs for the moderate group in the main analysis. Again, data points with RTs that fell above or below 2 *SDs* for each participant in the moderate AD group were omitted from the analysis.

Overall reaction time was analyzed using a 3 (Group: Mod AD vs Mild AD vs OAC) x 4 (item type: critical lure, unstudied unrelated critical lures, foils, and list items) mixed-model ANOVA was computed on correctly recognized items (and *probable* recognized items for the Moderate AD group). There was a significant main effect of item type, F (3, 87) = 338.54, $p < .001$, $\eta^2_p = .79$. Tukey's post-hoc tests showed greater RTs for CLs (7024 ms) compared to list items (5831 ms), foils (4030 ms), and unrelated CLs (4129 ms). Furthermore, RTs for list items were reliably slower than for foils and unrelated CLs. There were no reliable differences in RT for foils and unrelated CLs. As expected there was also a main effect of group, F (2, 87) = 185, $p < .001$, $\eta^2_p = .81$. Tukey's post-hoc tests revealed that the moderate AD group took reliably longer on the recognition task (6119 ms) compared to the mild AD group (5428.17 ms) and OACs (4214 ms) with the latter two also differing reliably. There was no significant interaction (*p* ns).

Overall, these data show that in line with the monitoring-process model, all groups spent longer on the critical lure words than the other item types. Importantly, and hitherto unreported, this pattern held for the moderate AD group (see Table 4). This latter finding is critical because it may indicate that although adults with moderate AD responded "new" to all items, there may have been some residual memory activation for presented items and CLs that caused the delay in rejecting these items as "new," with false memories having higher residual activation rates than list items (as predicted by associative-activation models). We now turn to whether these *probable* false memories that were *implicitly* activated can facilitate subsequent problem-solving.

INSERT TABLE 4 ABOUT HERE

CRAT Task

The mean CRAT solution rates (proportions) and the mean CRAT solution times (in seconds) were calculated for each participant and analyzed separately in a series of 3 (Group: Moderate AD vs. Mild AD vs. OAC) x 3 (Priming: primed/FM vs. primed/no-FM vs. unprimed) ANOVAs. For primed CRAT problems, solution rates and solution times were conditionalized on whether participants had produced the critical lure during recognition (OACs and mild AD) or by using reaction times for "new" recognition responses as a proxy for the moderate AD group creating three categories of priming (i.e., primed/FM = relevant DRM list was presented and the critical lure was falsely recognized and primed/no-FM = relevant DRM list was presented but the critical lure was not falsely recognized; note that priming effects are only found when CLs have been falsely recognized). The average percentage of primed items where the critical lure was falsely recognized (primed/FM) and was not recognized (primed/no-FM) are shown in Table 5. Thus, both solution rates and solution times were subjected to separate ANOVAs where the factors were solution type (unprimed, primed/no-FM, or primed/FM) and group. The data are shown in Table 6.

CRAT Solution Times. Concerning solution times, there was a main effect for priming $F(2, 87) = 83.89$, $p < .001$, $\eta^2_p = .5$, where Tukey's post-hoc tests showed that solution times were faster for primed/FM problems ($M = 27.59$ s) compared to primed/No-FM problems ($M = 37.5$ s, $p < .01$), and unprimed CRAT problems ($M = 37.16$ s, $p = < .01$), and the latter two conditions did not differ. There was also a main effect of group, $F(2, 87) = 134.08$, $p < .001$, $\eta^2_p = .76$. Tukey's post-hoc tests showed that the OACs were reliably faster at solving CRAT problems ($M = 27.17$ s) compared to Mild ($M = 33.52$ s) and Moderate (*M*

= 41.57 s) AD groups, with the latter two also differing reliably. There was no interaction (*p* ns).

CRAT Solution Rates. Concerning solution rates, there was a main effect for priming $F(2, 87) = 50.13$, $p < .001$, $\eta^2_p = .374$, where Tukey's post-hoc tests showed that solution rates were higher for primed/FM CRAT problems (M = 2.75) than for primed/no-FM ($M = 1.63$) and when participants were unprimed ($M = 1.60$), and the latter two did not differ. There was also a main effect for group, $F(2, 87) = 29.03$, $p < .001$, $\eta^2_p = .409$. Tukey's post-hoc tests revealed that the OACs solved reliably more CRATS ($M = 2.62$) compared to the Mild ($M = 2.00$) and Moderate ($M = 1.38$) AD groups, with the latter two groups also differing significantly. There was no interaction (*p* ns).

INSERT TABLES 5 AND 6 ABOUT HERE

What these findings tell us is that even though people with moderate AD do not consciously discriminate old from new items on recognition tests, saying "new" to everything, the time they take to say "new" differs as a function of item type. That is, just like OACs and people with mild AD, those with moderate AD take longer to respond to CL probes than items that were presented. Moreover, the CLs that were automatically activated in memory during encoding served as effective primes for all groups when it came to solve subsequently presented CRAT problems. Thus, it would seem that even for those people whose AD has progressed to a moderate level, associative networks may still be intact, and the information contained within them can be used to prime problem solutions.

Experiment 2

Having established the existence of false memory priming effects even in people with moderate AD (at least using CRAT problems), we wanted to extend these findings to more difficult problems (verbal proportional analogies) as well as examine whether primes are as effective when they are actually presented in DRM lists (true memories) as when they are self-generated (false memories). Thus, in our second experiment, we examined the propensity of healthy older adults, people with mild AD, and people with moderate AD to remember DRM lists and (automatically) self-generate CLs from those lists as well as their ability to use those false memories to aid their performance on subsequent analogical problem-solving tasks.

Method

A new sample of 90 participants was recruited whose demographic and other characteristics are shown in Table 1. Like Experiment 1, this study was also approved by the ethics committee at City, University of London (ethics approval reference PSYETH (T/F) 16/17 209).

Sixty participants had a clinical diagnosis of probable or possible AD (McKhann et al., 1984). Of these 60 participants with AD, 30 had mild AD and 30 had moderate AD. Participation took place in the participant's home or at a local Day Centre attended by the participant. The remaining 30 participants made up the OAC group. These people were community dwelling and were recruited from a panel of older adults who had expressed an interest in participating in research. Participation took place at City, University of London. There were no significant differences between OAC's and AD participants on age, years of education, and scores on the National Adult Reading Test, but as shown in Table 1, all three groups differed on most cognitive tests.

Design. A 3(Group: Moderate AD vs. Mild AD vs. OAC) x 5(Priming: true memory prime vs. true memory no priming vs. false memory prime vs. false memory no priming vs.

unprimed) mixed design was used. Group was a between-participants factor and priming was a within-participant factor. Primed problems were analyzed for those participants who either correctly remembered the critical lure when it was actually presented in the word list (true memory priming) or falsely remembered the critical lure when it was not presented in the word list (false memory priming). Finally, there were unprimed problems where no word list relevant to the problem had been presented.

Materials and Procedure. The procedure was essentially a replication of Akhtar and Howe (2019). All of the DRM lists and analogies used were previously normed (Akhtar & Howe, 2019) and are shown in Appendix B. The procedure had three stages: study phase, recognition task, and problem-solving task. There were nine DRM lists selected from Roediger et al. (2001). Each DRM list consisted of 12 words (e.g., *table, legs, seat, couch*) and was associated with an unpresented CL item (i.e., *chair*). Associative words that overlapped with the items in the analogical problems were removed and replaced with another associate.

Study phase. Participants were tested individually and were presented verbally with six of the nine DRM lists in a randomized order. Three of the lists contained the true memory prime (i.e., the first presented item in the list was the critical lure and analogy solution) and three of the lists did not contain the critical lure to three of the subsequently presented analogies (i.e., the critical lure/analogy solution would be the associated but unpresented critical lure to each of the three lists). Each list was presented verbally by the experimenter and was followed by a brief distractor task (counting backwards in threes). Following this, the next list was presented. This pattern of study then distractor task continued until all six lists were presented.

Recognition task. Following a five-minute distractor task, participants took part in a recognition task. Participants were verbally presented, in random order, with the 6 critical lure words (3 true and 3 false words) from the studied DRM lists, 6 unstudied and unrelated critical lure words (these were critical lures taken from unrelated and non-presented DRM lists), 32 list items (these were studied words from the presented DRM lists), and 32 unrelated words (these were words taken from unrelated and non-presented DRM lists; unrelated foils). All items in the recognition test were taken from Roediger et al. (2001). Again, we used a recognition test rather than a recall test to reduce effects of priming during retrieval (Olszewska & Ulatowska, 2013). For each word presented in the recognition task, participants had to select either [O], indicating that the word was Old and that they recognize the word from the previously presented lists, or [N] if they thought the current word presented was a new word that they did not hear in the previous word lists.

Analogical reasoning task. The final task was to solve verbal analogical-reasoning problems. Participants solved a practice problem before completing nine test analogical-reasoning problems. In a randomized order, participants were presented with nine analogical-reasoning problems in the format of '*a* is to *b* as *c* is to ?' For example, the analogy for the critical lure 'foot' was '*hat* is to *head* as *sock* is to _____.' Participants provided their answers verbally to the experimenter (participants first response was recorded). The time taken for them to complete the analogical problem, from presentation of the analogical problem to production of the response, was recorded. Participants were given a maximum of 60 s to provide an answer. Presentation of the DRM lists according to their link to the solution type in the unprimed, true memory prime, and false memory prime conditions was fully counterbalanced such that each DRM list and associated analogical-reasoning problem appeared equally often across participants within each group in each solution-type condition.

Results and Discussion

Recognition Task

Like in Experiment 1, we were interested in people's recognition rates and reaction times. Again, during the recognition task the moderate AD group responded 'N' (new) to all items (without exception) showing that they were not able to discriminate between the different item types. As such, the data from this group were removed from the recognition task analysis. The mean correct recognition rate (percentage) and the mean recognition reaction time (milliseconds) were calculated for each participant and analyzed separately in a series of 2(Group: Mild AD vs. OAC) x 4 (Item Type: Critical lures vs. Unrelated Critical Lures vs. List Items vs. Foils).

Recognition Rates. The recognition task showed that both the OACs and the Mild AD group created false memories for the critical lure words, with the Mild AD group falsely recognizing the critical lure 61.93% (M = 1.88, SD = 0.45) of the time and the OAC group 60% (M = 1.25, SD = .81) of the time (see Table 7). There were no reliable group differences (t ns).

When an ANOVA was conducted, there was a significant main effect of item type, $F(3, 57) = 857.31$, $p < .001$, $\eta^2_p = .93$. Pairwise comparisons revealed greater recognition for list items (77.22%) compared to foil items [correctly identifying these as 'New' items, (60.8%) (M = 24.71, SD = 3.16 vs M = 19.47, SD = 5.86)] and greater recognition of CL words (60%) compared to unrelated CL words (48.32%) (M = 4.2, SD = 1.18 vs. M = 2.57, SD = 1.22). There was also a main effect of group $F(1, 58) = 42.44$, $p < .001$. To investigate this further, we conducted separate t-tests on item types; for list items, the OAC group recognized significantly more items (M = 26.32, SD = 3.16) than people with mild AD (M = 23.1, SD = 3.15) t (59) = 3.5, $p < .001$. Concerning foil items, again the OAC group correctly rejected significantly more items (*M* = 22.62, *SD* = 6.2) than people with mild AD (*M* = 16.31, *SD* = 3.3), t (59) = 2.28, $p < .05$.

INSERT TABLE 7 ABOUT HERE

Signal Detection Analysis. We again ran signal detection analysis. The values of *d'* and *C* are shown in Table 8. Signal detection measures for hits and critical lures were analyzed in separate 2 (Group: Mild AD vs OAC) x 2 (Item Type: critical lures vs studied items) mixed-model ANOVA. The analysis of *d*' revealed a main effect of item type, where discriminability was better for critical lures compared to list items, $F(1, 58) = 46.67$, $p < .001$, $\eta^2_p = .17$. There was no main effect of group and no interaction (*p*s ns). Analysis of the criterion C revealed a more liberal bias for the critical lures $F(1, 58) = 7.2$, $p < .001$, $\eta^2_p = .35$ compared to studied items. There was also a main effect of group $F(1, 58) = 18.29$, $p < .001$, $\eta^2_p = .41$, where people with Mild AD revealed a more liberal bias compared to OACs. There was no significant interaction.

INSERT TABLE 8 ABOUT HERE

Recognition Times. Again, like in Experiment 1, we were interested in looking to see if the moderate AD group automatically created *probable* CLs (based on their latencies). We followed the same methodology as Experiment 1, by first comparing OACs and mild AD group reaction times. RTs that fell below 2 *SDs* were removed from the analysis.

We ran a 2 (Group: OAC vs. Mild AD) x 4 (Item type: CLs vs. List Items vs. Unrelated unpresented CLs vs. foils) mixed-model ANOVA (see Table 9). There was a main effect of item type [$F(3, 57) = 1095.5$ $p < .001$, $\eta^2_p = .95$]. Tukey's post-hoc tests indicated longer RTs for CLs (6458 ms), followed by list items (5008 ms) [p <.001], foils (3715 ms) and unrelated CLs (3677 ms) (there were no reliable difference between the latter two). As expected, there was a main effect of group $F(1, 58) = 617.8$, $p < .001$, $\eta^2_p = .91$. Tukey's post-hoc tests showed the mild AD group were reliably slower (5264.5 ms) compared to the OACs (4165 ms). There was no interaction $p > .05$. Like in Experiment 1 we ran separate analyses comparing CLs vs correct rejections [$F(1, 58) = 647.17$ $p < .001$, $\eta^2_p = .92$] and found that false recognition of critical lures was reliably slower (6458 ms) compared to correct rejections of CLs (5046 ms). We also compared correct recognition of list items to those items that were not correctly recognized [$F(1, 58) = 609$, $p < .001$, $\eta^2_p = .91$] and found correctly recognized list items were reliably slower (5008 ms) than those list items that were not correct recognized (3703 ms). There was no difference in reaction time for correctly rejected foils compared to foils not rejected ($p > .05$) and no difference in reaction time between correctly rejected unrelated CLs versus nonrejected unrelated CLs. Next, just like in Experiment 1, we compared these latency pattern to those from the moderate AD group.

Again, data points with RTs that fell above or below 2 *SDs* for each participant were omitted from the analysis. There was a significant main effect of item type, $F(3, 86) = 1129.62$, $p < .001$, $\eta^2_p = .93$. Tukey's post-hoc tests showed longer RTs for CLs ($M = 6973$ ms, $p = < .01$) compared to list items ($M = 5465$ ms, $p = < .01$), foils ($M = 3856$ ms), and unrelated CLs ($M = 3839$ ms). There were no reliable differences between the latter two. There was also a main effect of group, $F(2, 87) = 507$, $p < .001$, $\eta^2_p = .9$. Tukey's post-hoc tests revealed that the Moderate AD group took reliably longer on the recognition task ($M = 5671$ ms, $p < .01$) compared to the Mild AD group ($M = 5264$ ms, $p < .01$), and the OACs ($M = 4165$ ms, $p < .01$). There was no interaction (p ns).

Like in Experiment 1, these data show that all three groups spent longer on the critical lure words than the other item types. Again, this pattern was the same for the Moderate AD group (see Table 9). This is a critical finding inasmuch as it may indicate that although adults with moderate AD responded "new" to all items, there may have been some residual memory activation for presented items and CLs that caused the delay in rejecting these items as new, with false memories having higher residual activation rates than list items (as predicted by associative-activation models).

INSERT TABLE 9 ABOUT HERE

Analogical Reasoning Task

The mean analogy solution rates (proportions) and the mean analogy solution times (in seconds) were calculated for each participant and analyzed separately in a series of 3(Group: OAC vs. Mild AD vs. Moderate AD) x 5(Priming: primed/FM vs. primed/no-FM vs. primed/TM vs. primed/no-TM vs. unprimed) ANOVAs. Like in Experiment 1, problem solving data were conditionalized on memory performance and we used the same reaction time proxy measure for the moderate AD group. Again, primed/FM refers to analogy problems that were solved correctly only when the related false memory was recognized on the memory test and primed/no-FM refers to problems solved when the DRM list was presented but the participant did not falsely remember the critical lure on the recognition test. Primed/TM refers to analogy problems that were solved correctly only when the true memory was recognized on the memory test and primed/no-TM refers to problems solved when the DRM list was presented but the participant did not remember the actually presented critical lure on the recognition test. Of course, unprimed refers to analogy problems that were solved correctly when no DRM priming list was presented. The average percentage of primed items where the critical lure was falsely (primed/FM) or correctly (primed/TM) recognized and was not falsely (primed/no-FM) or correctly (primed/no-TM) recognized are shown in Table 10.

Analogy Solution Times. There were significant differences in solution times as a function of group and priming. Specifically there was a main effect for priming, $F(4, 87) = 105.45$, $p <$ .001, $\eta^2_p = .551$, where Tukey's post-hoc tests showed that solution times were faster for primed/FM problems (*M* = 4.33 s) than for primed/no-FM problems (*M* = 6.54 s, $p < .01$), primed/TM problems (*M* = 6.59 s, $p < .01$), primed/no-TM problems (*M* = 6.69 s, $p < .01$), and unprimed problems (*M* = 6.68 s, *p* = < .01), and the latter four conditions did not differ. There was also a main effect of group, $F(2, 87) = 555.67$, $p < .001$, $\eta^2_p = .925$, where Tukey's post-hoc tests showed OAC's solution times were faster (*M* = 4.28 s) compared to people with Mild AD (*M* = 6.09 s) and Moderate AD (*M* = 8.13 s), with the latter two groups also differing significantly. There was no interaction (*p* ns; see Table 11).

INSERT TABLES 10 AND 11 ABOUT HERE

Analogy Solution Rates. Concerning solution rates, there was a main effect for priming, $F(4, 87) = 75.83$, $p < .001$, $\eta^2_p = .466$, where Tukey's post-hoc tests showed that solution rates were higher for primed/FM analogical problems (*M* = 2.22) than for primed/no-FM (*M* = 0.57, $p < .01$), primed/TM (*M* = 1.56, $p < .01$), primed/no-TM (*M* = 1.06, $p < .01$), and when participants were unprimed (*M* = 1.52, $p < .01$), with the latter 3 conditions not differing. As expected, there was also a main effect of group, $F(2, 87) = 533.84$, $p < .001$, $\eta^2_p = .925$, where Tukey's post-hoc tests showed the OACs solved reliably more problems compared to people with Mild AD and Moderate AD, with the latter two groups also differing reliably (see Table 11). There was no significant interaction.

These results, like those from Experiment 1, clearly indicate that despite the fact that people with moderate AD fail to consciously discriminate old from new items on recognition tests (again, saying "new" to everything), the time they take to say "new" differs as a function of item type. In fact, like OACs and people with mild AD, people with moderate AD take longer to respond to CL probes than items that were actually presented. These CLs also served as effective primes regardless of group when it came time to solve subsequently presented analogical-reasoning problems. Moreover, like the OACs and the mild AD participants, problem-solving for people with moderate AD was facilitated only by CLs that were self-generated (i.e., false memories) and not by CLs that had actually been presented (i.e., true memories). Therefore, it seems clear that even for those people whose AD has progressed to a moderate level, associative activation within memory networks occurs automatically and the resultant activation can be used to prime analogical problem solutions.

General Discussion

The present study set out to extend the positive consequences of false memories using CRAT problems (Experiment 1) and analogical-reasoning tasks (Experiment 2). Although we showed the positive consequences of false memories in people with mild AD and OACs in Akhtar et al. (in press) and Akhtar and Howe (2019), here we not only replicated these findings but also extended them to people with moderate Alzheimer's Disease. Across both experiments using two different types of problem-solving tasks, we found that all three groups' performance on these tasks was facilitated by self-generated false memories that arose from the prior presentation of DRM lists. That is, performance on CRATs (Experiment 1) and analogical-reasoning tasks (Experiment 2) was better (greater accuracy and speed) when those problems were preceded by DRM lists whose CLs also served as the solution to those problems.

Our experiments show that like people with mild AD, those classified as having moderate AD also create false memories upon presentation of DRM lists. That is, although we found that in both experiments the moderate AD group responded "new" to all items, the pattern of response times was similar across the different item types for all participants regardless of group. Specifically, although participants with moderate AD took longer to respond overall than those with mild AD who, in turn, took longer overall to respond than healthy older adult controls, like the latter two groups, those with moderate AD took longer to reject CLs than actual list items.

This finding is important because it demonstrates that there are other, long-term priming techniques that can be used to demonstrate the preservation of implicit memory in people experiencing moderate AD. Equally important, these priming effects are conceptual in nature. That is, they are based on the creation of false memories that arise from the activation of semantically-related information contained in associative networks. Thus, consistent with the literature concerning the decline of explicit memory, but the sparing of implicit memory, in people with AD (e.g., see Fleishman et al., 2005; Meiran & Jelicic, 1995), our results show that associative memory networks are still intact in people with moderate AD and that spreading activation mechanisms are still operating within these networks.

Importantly, we also found that OACs and the mild AD group took reliably longer to falsely recognize CLs compared to the time it took them to correctly reject CLs. In addition, both groups were reliably slower to correctly respond “old” to list items compared to those that were rejected as “new.” When analyzing the latencies for the moderate AD group, we again found that their response times followed the same pattern as the older adults and the mild AD group. That is, although they did not explicitly recognize false memories on the recognition test, their response latencies to reject items suggested that some type of *implicit* monitoring for these items was taking place. We concluded from this that the moderate AD group had created *probable* false memories. Critically, like the OACs and the mild AD group, these *probable* false memories created by those with moderate AD facilitated subsequent problem-solving performance, further vindicating the use of recognition response times to diagnose the contents of memory. Together, these data show that although people with moderate AD were unable to consciously or explicitly acknowledge the presence of (true and) false memories, using response times, we could see that such memories had been generated during DRM list presentation.

Importantly, like previous research with children and adults (e.g., Howe et al., 2015), priming only occurred when CLs were self-generated and not when they were actually presented as part of the studied list (Experiment 2). As mentioned earlier, false memories are not simply false positive errors on a memory test. Rather, they are the result of the activation of one's own semantic-associative memory networks and in that sense, they are a self-generated memory response. Based on our results, it would appear that even individuals with moderate AD have intact memory networks whose activation can result in self-generated false memories, memories that can serve as better primes for subsequent problem-solving than other-generated information that was actually presented during study ("true" memories). Thus, the data reported here demonstrates that self-generated information from semantic networks is important for memory not just in children (e.g., Cunningham et al., 2014; Ford & Lobao, 2018), young adults (e.g., Mulligan & Lozito, 2004), old adults (e.g., Gutchess & Kensinger, 2018), and old adults with mild AD (Klein et al., 2003; Rosa et al., 2015), but now also extends to older adults with moderate AD.

To summarize, our findings support and extend current research efforts to understand what aspects of memory remain intact in people with mild to moderate AD. Specifically, our results extend previous research on false memory formation in people with mild AD (e.g., Akhtar & Howe, 2019; Akhtar et al., in press; Evrard et al., 2016; Malone et al., 2019) to people experiencing moderate AD. Our findings also support the use of reaction time data when understanding the contents of memory, particularly in those populations whose (recognition) memory responses do not necessarily reflect what is actually in their memory. This is especially true when some contents of memory, in particular the creation of false memories, are derived automatically through semantic-associative activation.

Our findings also illustrate the importance of self-generated information in memory. Indeed, Experiment 2 showed that self-generated (false) memories facilitated subsequent problem-solving but those same memories when actually presented during study to participants had no effect on subsequent problem-solving. This is consistent with previous research with children and young adults (e.g., Howe et al., 2015) as well as work concerning the importance of self-referencing effects in false memory formation in people with mild AD (e.g., Rosa et al., 2015). Our data extend these findings to include people whose AD has progressed from mild to moderate levels. Importantly, then, even people with moderate AD automatically form self-generated (false) memories and those memories are more likely to prime solutions to subsequently presented problems than those memories that are based on information that was actually studied.

Finally, our findings are consistent with and extend previous research on semantic priming effects in people with AD. As Gilet et al. (2017) discovered, although people with AD produced fewer CLs on a memory test than OACs, they, like the healthy controls, exhibited shorter lexical decision latencies for items that were preceded by a DRM list whose CL was related to the lexical decision item. Thus, like our findings, even though some participants were not explicitly aware of the contents of their memory, automatically formed false memories could be used to facilitate performance on a subsequent task. Our experiments extend these findings in two ways. First, we demonstrated that priming effects occur not just in lexical decision tasks but also when solving more complex CRAT and analogical-reasoning problems. Second, these priming effects are not only present in people with mild AD but also in those diagnosed with moderate levels of AD. Importantly, although not explicitly recognized as being in memory, by using analyses of response times to items on a recognition test, we could discover that false memories had been formed and that those memories could prime subsequent problem-solving solutions. Perhaps by using a more sensitive, implicit measure of memory (response times), we can discover the contents of memory in people who cannot explicitly acknowledge those contents on a memory test. Moreover, by providing a subsequent (e.g., problem-solving) task where performance advantages can arise only if specific items are in memory, we provide a second, independent measure of the contents of memory. Thus, when these measures are in place, it seems that even people with moderate AD can and do form false memories and those self-generated memories can prime and facilitate performance on subsequent problem-solving tasks.

References

Akhtar, S., & Howe, M. L. (2019). Priming older adults and people with Alzheimer's disease analogical problem-solving with true and false memories. *Journal of Clinical and Experimental Neuropsychology*, *41*, 704-714.

Akhtar, S., Howe, M. L., & Hoepstine, K. (in press). Can false memories prime problem solutions for healthy older adults and those with Alzheimer's disease? *Journal of Gerontology: Psychological Sciences*.

Akhtar, S., Moulin, C. J., & Bowie, P. C. (2006). Are people with mild cognitive impairment aware of the benefits of errorless learning? *Neuropsychological Rehabilitation*, *16*, 329-346.

Atkinson, R. C., & Juola, J. F. (1974). *Search and decision processes in recognition memory*. WH Freeman.

Coane, J. H., McBride, D. M., Raulerson III, B. A., & Jordan, J. S. (2007). False memory in a short-term memory task. *Experimental Psychology, 54*, 62-70.

Cunningham, S. J., Brebner, J. L., Quinn, F., & Turk, D. J. (2014). The self-reference effect on memory in early childhood. *Child Development, 85*, 808-823.

Deese, J. (1959). On the prediction of occurrence of particular verbal intrusions in immediate recall. *Journal of Experimental Psychology*, *58*, 17-22.

Dew, I. T. Z., & Giovanello, K. S. (2010). Differential age effects for implicit and explicit conceptual associative memory. *Psychology and Aging, 25*, 911-921.

Evrard, C., Colombel, F., Gilet, A.-L., & Corson, Y. (2016). Intact semantic priming of critical lures in Alzheimer's disease: Implications for false memory. *Journal of Gerontology: Psychological Sciences, 71,* 671-674.

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*, 175-191.

Fleischman, D. A., Wilson, R. S., Gabrieli, J. D. E., Schneider, J. A., Bienias, J. L., & Bennett, D. A. (2005). Implicit memory and Alzheimer's disease neuropathology. *Brain, 128*, 2006-2015.

Ford, R., & Lobao, S. N. (2019). Exploring individual differences in self-reference effects for agency and ownership in 5- to 7-year-olds. *British Journal of Developmental Psychology, 37*, 168-183.

Gilet, A.-L., Evrard, C., Colombel, F., Tropée, E., Marie, C., & Corson, Y. (2017). False memories in Alzheimer's disease: Intact semantic priming but impaired production of critical lures. *Journal of Gerontology: Psychological Sciences, 72,* 986-990.

Gutchess, A., & Kensinger, E. A. (2018). Shared mechanisms may support mnemonic benefits from self-referencing and emotion. *Trends in Cognitive Sciences, 22*, 712-724.

Hodges, J. R., Salmon, D. P., & Butters, N. (1992). Semantic memory impairment in Alzheimer's disease: Failure of access or degraded knowledge? *Neuropsychologia*, *30*, 301-314.

Howard, D., & Patterson, K. E. (1992). *The Pyramids and Palm Trees Test: A test of semantic access from words and pictures*. Thames Valley Test Company.

Howe, M. L., Garner, S. R., Charlesworth, M., & Knott, L. M. (2011). A brighter side to memory illusions: False memories prime children's and adults' insight-based problem solving. *Journal of Experimental Child Psychology*, *108*, 383-393.

Howe, M. L., Garner, S. R., Threadgold, E., & Ball, L. J. (2015). Priming analogical reasoning with false memories. *Memory & Cognition*, *43*, 879-895.

Howe, M. L., Threadgold, E., Norbury, J. V., Garner, S. R., & Ball, L. J. (2013). Priming children's and adults' analogical problem solutions with true and false memories. *Journal of Experimental Child Psychology*, *116*, 96-103.

Howe, M. L., Wilkinson, S., Garner, S. R., & Ball, L. J. (2016). On the adaptive function of children's and adults' false memories. *Memory, 24*, 1062-1077.

Howe, M. L., Wilkinson, S., & Monaghan, P. (2012). *False memories trump true ones as problem-solving primes after a delay.* Paper presented at the annual meeting of the Psychonomic Society, Minneapolis, MN.

Howe, M. L., Wimmer, M. C., Gagnon, N., & Plumpton, S. (2009). An associative-activation theory of children's and adults' memory illusions. *Journal of Memory and Language*, *60*, 229-251.

Humphreys, G. W., & Sui, J. (2016). Attentional control and the self: The self-attention network. *Cognitive Neuroscience, 7*, 5-29.

Jou, J., Matus, Y. E., Aldridge, J. W., Rogers, D. M., & Zimmerman, R. L. (2004). How similar is false recognition to veridical recognition objectively and subjectively? *Memory & Cognition*, *32*, 824-840.

Klein, S. B., Cosmides, L., & Costabile, K. A. (2003). Preserved knowledge of self in a case of Alzheimer's dementia. *Social Cognition, 21*, 157-165.

Kučera, H., & Francis, W. N. (1967). *Computational analysis of present-day American English*. Dartmouth Publishing Group.

Malone, C., Deason, R. G., Palumbo, R., Heyworth, N., Tat, M., & Budson, A. E. (2019). False memories in patients with mild cognitive impairment and mild Alzheimer's disease dementia: Can cognitive strategies help? *Journal of Clinical and Experimental Neuropsychology, 41,* 204-218.

McKhann, G., Drachman, D., Folstein, M., Katzman, R., Price, D., & Stadlan, E. M. (1984). Clinical diagnosis of Alzheimer's disease: Report of the NINCDS- ADRDA Work Group* under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease. *Neurology*, *34*, 939-939.

Mednick, S. (1962). The associative basis of the creative process. *Psychological Review, 69,* 220-232.

Meiran, N., & Jelicic, M. (1995). Implicit memory in Alzheimer's disease: A meta-analysis. *Neuropsychology, 9*, 291-303.

Mulligan, N. W., & Lozito, J. P. (2004). Self-generation and memory. *Psychology of Learning and Motivation, 45*, 175-214.

Naveh-Benjamin, M., Shing, Y. L., Kilb, A., Werkle-Bergner, M., & Lindenberger, U. (2009). Adult age differences in memory for name-face associations: The effects of intentional and incidental learning. *Memory, 17*, 220-232.

Nebes, R. D. (1989). Semantic memory in Alzheimer's disease. *Psychological Bulletin, 106,* 377-394.

Olszewska, J., & Ulatowska, J. (2013). Encoding strategy affects false recall and recognition: Evidence from categorical study material. *Advances in cognitive psychology*, *9*, 44-52.

Roediger III, H. L., Balota, D. A., & Watson, J. M. (2001). Spreading activation and arousal of false memories. In H. L. Roediger, J. S. Nairne, I. Neath, & A. M. Surprenant (Eds.), *The nature of remembering: Essays in honor of Robert G. Crowder* (pp. 95-115). Washington, DC: American Psychological Association.

Roediger, H. L., & McDermott, K. B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*, 803-814.

Roediger, H. L., 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.

Rosa, N. M., Deason, R. G., Budson, A. E., & Gutchess, A. H. (2015). Self-referencing and false memory in mild cognitive impairment due to Alzheimer's disease. *Neuropsychology*, *29*, 799-805.

Tun, P. A., Wingfield, A., Rosen, M. J., & Blanchard, L. (1998). Response latencies for false memories: Gist-based processes in normal aging. *Psychology and Aging*, *13*, 230-241.

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

Sui, J., & Humpreys, G. W. (2015). The integrative self: How self-reference integrates perception and memory. *Trends in Cognitive Sciences, 19*, 719-728.

Table 1. Means (and Standard Error) Demographic Characteristics of Participants

	Moderate AD	Mild AD	OAC	P Value
Age	73.54 (7.2)	74.55 (7.1)	71.24 (7.8)	NS
Education (years)	10.9	11.8	12.1	NS
NART estimated IQ	115.32 (8.7)	116.28 (6.9)	114.4 (7.8)	NS
MMSE (out of 30)	18.9 (2.8)	25.8 (2.2)	29.8 (0.5)	<0.0001*§+
Range (defines minimal and mild)	17-20	24-27	29-30	
CERAD				
Immediate (max = 30)	2.6 (2.44)	13.1 (2.36)	17.8 (3.96)	<0.0001*§+
Delayed (out of 10)	0	1.9 (1.97)	6.8 (3.58)	<0.0001*
Recognition (out of 10)	0	5.9 (3.1)	9.4 (3.4)	<0.0001*
Digit span (out of 10)				
FWD	2.9 (4.3)	5.3 (1.2)	6.9 (1.87)	<0.0001*§+
BCK	0	2.99 (2.1)	5.6 (1.54)	<0.0001*

*Significant difference between OAC and Mild AD
§ Significant difference between OAC and Moderate AD
+ Significant difference between Mild AD and Moderate AD
NS, difference not significant

Table 2. Percentage of correct items recognized in the recognition task for OACs and people with Mild AD.

	Critical Lures	List Items	Foils	Unrelated CL
OAC	63.01% (1.23)	85.26% (3.2)	70.09% (4.59)	66.23% (3.99)
Early AD	64.93% (6.31)	74.4% (1.97)	50.9% (2.9)	49.57% (2.11)
Moderate AD*	62.4% (3.23)	49.26% (2.9)	-	-

**probable* FMs and *probable* List Items based on latencies

Table 3. Means and Standard deviations of Signal Detection Measures of Discriminability (*d'*) and Bias (*C*) for Studied Items and Critical Lures (CL).

	OACs		Mild AD	
	d'	*C*	*d'*	*C*
Studied	0.83 (0.46)	0.89 (0.41)	0.35 (0.6)	0.51 (0.22)
Lures	0.79 (0.54)	0.27 (0.26)	1.02 (0.11)	0.16 (21.4)

Table 4. Reaction time for the recognition task for OACs, Mild AD, and Moderate AD group.

	Critical Lures	Correct rejections CLs	Test statistic *t*	List Items	Unrecognized List Items	Test statistic *t*	Foils	Unrelated CL
Moderate AD	8326 (438)	6593 (353)	17.2**	6898 (666)	4969 (592	11.94**	4617 (578)	4636 (604)
Mild AD	7106 (409)	5672 (287)	17.7**	5881 (806)	4273 (637)	12.85**	4283 (736)	4441 (827)
OAC	5640 (619)	4626 (504)	9.8**	4716 (563)	3160 (421)	14.5**	3191 (607)	3310 (1041)

** Significant p <.001

Table 5. Percentage of CRATs that were primed by false memories where the critical lure was recognized and where it was not recognized on the prior memory test.

	Primed/FM	Primed/No-FM
OACs	63.01	36.99
Mild AD	64.93	35.07
Moderate AD*	62.14	37.86

*_probable_ percentages based on latencies

Table 6. Mean CRAT problem solution times and solution rates for older adults, people with Mild AD, and people with Moderate AD as a function of priming.

Participant	Priming		
	Unprimed	Priming/FM	Priming/NO-FM
Solution times (seconds)			
Older Adults	30.53 (4.54)	20.4 (4.3)	30.57 (4.44)
Mild AD	37.46 (4.37)	26.57 (6.4)	36.54 (3.24)
Moderate AD	43.48 (9.01)	35.79 (5.3)	45.39 (5.28)
Solution rates (proportion)			
Older Adults	0.42 (0.94)	0.67 (0.86)	0.45 (0.19)
Mild AD	0.31 (1.02)	0.56 (0.68)	0.31 (0.91)
Moderate AD	0.20 (0.68)	0.40 (0.74)	0.21 (0.961)

Note: Standard errors are in parenthesis. FM = False Memory

Table 7. Percentage of correct items recognized in the recognition task for OACs and Mild AD group.

	Critical Lures	List Items	Foils	Unrelated CL
OAC	60.01% (1.89)	82.26% (1.89)	70.69% (6.3)	64.89% (5.69)
Mild AD	61.93% (2.97)	72.19% (2.27)	51% (3.28)	44.39% (6.45)
Moderate AD*	62.18% (2.4)	48.34% (2.57)	-	-

**probable* FMs and List Items based on latencies

Table 8. Means and Standard deviations of Signal Detection Measures of Discriminability (*d'*) and Bias (*C*) for Studied Items and Critical Lures (CL).

	OACs		AD	
	d'	C	d'	C
Studied	0.81 (0.46)	0.89 (0.41)	0.35 (0.6)	0.51 (0.22)
Lures	0.97 (0.54)	0.24 (0.23)	1.05 (0.11)	0.16 (21.4)

Table 9. Reaction time for the recognition task for OACs, Mild AD, and Moderate AD group.

	Critical Lures	Correct rejections CLs	Test statistic *t*	List Items	Unrecognized List Items	Test statistic *t*	Foils	Unrelated CL
Moderate AD	8002(410)	6320 (403)	19.3**	6380 (499)	4266 (533)	16.92**	4139 (703)	4164 (535)
Mild AD	7181(430)	5540 (336)	17.29**	5543 (321)	4229 (293)	18.92**	4191 (311)	4141 (276)
OAC	5735 (281)	4552 (274)	20.57**	4473 (269)	3178 (316)	16.25**	3239 (304)	3212 (265)

Table 10. Percentage of analogies that were primed by true and false memories where the critical lure was recognized and where it was not recognized on the prior memory test.

	Primed/FM	Primed/No-FM	Primed/TM	Primed/No-TM
OACs	60.01	39.99	84.2	15.8
Mild AD	61.93	38.07	73.56	26.44
Moderate AD*	62.18	37.82	52.16	47.84

**probable* percentages based on latencies

Table 11. Mean analogical problem solution times and solution rates for older adults, people with Mild AD, and people with Moderate AD as a function of priming.

Participant	Priming				
	Unprimed	Primed/TM	Primed/no-TM	Primed/FM	Primed/no-FM
Solution times (seconds)					
OAC	4.71 (1.15)	4.63 (0.67)	4.84 (1.27)	2.56 (0.71)	4.66 (0.12)
Mild AD	6.59 (0.9)	6.48 (0.87)	6.4 (0.7)	4.5 (0.76)	6.48 (0.71)
Moderate AD	8.74 (1.41)	8.66 (1.27)	8.83 (1.02)	5.93 (1.05)	8.48 (0.93)
Solution rates (proportions)					
OAC	0.72 (0.69)	0.73 (0.71)	0.27 (0.71)	0.97 (.3)	0.03 (0.30)
Mild AD	0.5 (0.75)	0.52 (0.73)	0.48 (0.72)	0.73 (.76)	0.25 (0.72)
Moderate AD	0.3 (0.48)	0.31 (0.52)	0.31 (0.45)	0.52 (.72)	0.28 (0.46)

Note: Standard errors are in parenthesis. FM = False Memory, TM = True Memory

Appendix A. DRM and CRAT lists for Experiment 1.

DRM	CRAT
PEN - pencil, write, fountain, quill, felt, bic, scribble, cross, light, marker	PEN - pal, tip, knife
FOOT - walk, hand, toe, kick, sandals, yard, ankle, boot, inch, sock	FOOT - hold, print, stool
COLD - hot, snow, warm, winter, ice, chilly, heat, weather, freeze, frost	COLD - cut, sore, war
FRUIT - apple, vegetable, orange, citrus, ripe, banana, berry, cherry, basket, cocktail	FRUIT- bowl, Juice, Salad
CUP - mug, saucer, coaster, lid, handle, coffee, soup, drink, plastic, sip	CUP - measuring, cake, tea
WISH - want, dream, desire, Hope, well, think, star, ring, wash, true	WISH - list, death, bone
ARMY - navy, soldier, rifle, draft, military, marines, infantry, war, pilot, combat	ARMY - base, tank, territorial
LONG - short, narrow, John, time, far, hair, road, thin, line, low	LONG - bow, haul, jump
SHIRT- blouse, sleeves, Trousers, tie, button, shorts, collar pocket, jersey, cuffs	SHIRT - polo, flannel, vest
MOUNTAIN - hill, valley, climb, summit, molehill, peak, plain, climber, steep, ski	MOUNTAIN - bike, top, goat

Appendix B. Analogies and DRM lists for Experiment 2.

Water:Boat::Road:Car

Car – truck, bus, train, vehicle, drive, jeep, ford, race, keys, garage, highway, van

Stand:Floor::Sit:Chair

Chair – table, legs, seat, couch, desk, recliner, sofa, wood, cushion, swivel, stool, sitting

Tooth:Brush::Hair:Wash

Wash – rinse, dishes, mouth, scrub, laundry, soap, shampoo, dish, soak, cloth, bathroom

Desert:Hot::Arctic:Cold

Cold – snow, warm, winter, ice, wet, frigid, chilly, heat, weather, freeze, air, shiver

Poverty:Wealth::Sickness:Health

Health – sickness, good, happiness, ill, service, strong, disease, body, poorly, pain, vigor, robust

Tunnel:Mountain::Bridge:River

River – water, stream, lake, Mississippi, boat, tide, swim, flow, run, barrage, creek, brook

Fire:Hot::Candy:Sweet

Sweet – sour, sugar, bitter, good, taste, tooth, nice, honey, soda, chocolate, heart, cake

Dark:Light::Short:Long

Long – tall, narrow, John, time, far, hair, island, road, thin, underwear, distance, line

Pestle:Mortar::Saucer:Cup

Cup – mug, tea, measuring, coaster, lid, handle, coffee, straw, goblet, soup, stein, drink