Free Recall and Learning of Emotional Word Lists in Very Elderly People With and Without Dementia

Ruth E. Nieuwenhuis-Mark, Kim Schalk and Natalie de Graaf

Am J Alzheimers Dis Other Demen 2009; 24; 155 originally published online Jan 26, 2009;
DOI: 10.1177/1533317508330561

The online version of this article can be found at:
http://aja.sagepub.com/cgi/content/abstract/24/2/155

Published by:

SAGE
http://www.sagepublications.com

Additional services and information for American Journal of Alzheimer's Disease and Other Dementias® can be found at:

Email Alerts: http://aja.sagepub.com/cgi/alerts

Subscriptions: http://aja.sagepub.com/subscriptions

Reprints: http://www.sagepub.com/journalsReprints.nav

Permissions: http://www.sagepub.com/journalsPermissions.nav

Citations http://aja.sagepub.com/cgi/content/refs/24/2/155
Free Recall and Learning of Emotional Word Lists in Very Elderly People With and Without Dementia

Ruth E. Nieuwenhuis-Mark, PhD, Kim Schalk, MSc, and Natalie de Graaf, MSc

An emotional memory advantage has been found across the life span where recall is better for emotional (as opposed to neutral) stimuli. Our goal was to design emotionally valent word lists for easy use by practitioners and to test whether demented and healthy elderly participants showed an emotional memory advantage with these lists. Three new word lists (a positive, a negative, and a neutral list) were constructed. Thirty-eight controls, 37 with mild cognitive impairment and 20 Alzheimer’s dementia participants’ free recall was tested. Unsurprisingly, controls had better recall overall. Emotionally valent words were recalled better in comparison to neutral words in all 3 groups. No recall advantage for positive versus negative words emerged. Learning differed among the groups with the Alzheimer’s dementia participants showing flatter learning curves. The results tentatively suggest that emotional memory may stay intact longer but that learning of such lists becomes more difficult as dementia progresses.

Keywords: dementia; recall; learning; emotional word lists

Many types of dementia are characterized by memory impairment. However, as Fleming et al1 rightly pointed out, memory is not a unitary concept and different mnemonic functions depend on different neuroanatomical regions and/or networks. In relation to dementia, for example, several studies have suggested that (some) aspects of implicit memory stay intact long into the dementia process2,3 while other studies have suggested that memory for emotional stimuli may be to some extent spared1,4 but see ref 5. Alzheimer’s dementia (AD) is the commonest form of dementia and encoding of explicit declarative information is the primary memory deficit in AD. Such encoding depends on the integrity of the hippocampus and its surrounding regions.6 As dementia progresses, the hippocampus becomes more damaged and episodic memory in turn becomes more of a problem for these patients. The amygdala also degenerates in AD.7,8 Both the amygdala and the hippocampus are involved in, among other functions, memory for emotional stimuli/events.9 Emotional stimuli tend to be better remembered (ie, memory for these stimuli is both more accurate and more quickly processed) on the whole than neutral stimuli—the so-called emotional memory advantage (EMA). This effect has been observed for pictures,10 words,11 autobiographical experiences,12 and sounds.13 Whether there is a valence effect (negatively valenced stimuli better remembered than positive or vice versa) remains controversial. It is very likely that many cognitive and neural processes contribute to the EMA.9 Differences in encoding strategies are believed, by many researchers, to be one of the possible mechanisms behind the EMA. Emotional information may be more elaborately encoded, have a wider semantic network, be more easily linked with autobiographical experiences, and/or may be more distinctive than neutral items.9 The EMA could feasibly also be due to retrieval processes where emotion or mood may serve as a retrieval cue, or it may be due to modulation of consolidation processes. What the precise mechanism behind the EMA is still not known. Kensinger et al9 suggested that people who suffer from amygdala lesions do not show EMAs.
Following this logic to its conclusion, we may expect to see EMAs in people with mild AD and less atrophy of the amygdala in comparison to people with severe AD and greater atrophy of the amygdala. Kensinger et al.\textsuperscript{9} found that this was indeed the case when memory for emotional stories was tested.

Healthy elderly people also show an enhanced memory for emotional compared to neutral information.\textsuperscript{1,14,15} Although Kensinger and Corkin\textsuperscript{16} found that young participants recalled more negative stimuli in comparison to positive or neutral, several studies have reported that older adults show a disproportional enhancement of memory in favor of positive compared to negative stimuli. This phenomenon has been called the positivity effect.\textsuperscript{17,18} Kensinger\textsuperscript{19} went on to show that this positivity effect may only be obtained when nonarousing words are used. She found no valence differentiation when arousing words were used (arousing words were simply remembered better than neutral words and there was no positivity effect). Kensinger suggested that aging alters the processing of nonarousing information due to the engagement of controlled strategies while EMAs to arousing stimuli are more likely to be governed by more automatic processes (the latter are believed to be less affected by aging). According to Mather and Carstensen,\textsuperscript{20} the positivity effect could be explained by the fact that older adults are more likely to focus on positive information in everyday life, and less on negative information, in a strategy designed to enhance their sense of well-being. However, a number of studies have reported a memory enhancement for negative emotional stimuli in healthy elderly people as well.\textsuperscript{14,15}

Relatively few studies have been carried out in this research area in people with mild cognitive impairment (MCI) or with varying degrees of dementia. Petersen and colleagues\textsuperscript{21-23} see MCI as an entity in its own right. They have provided criteria to assist clinicians when they suspect “MCI but not yet dementia” in their patients. These authors have also suggested that subgroups of MCI exist with the amnesic form being the most likely to develop into full-blown AD. Marrreneca et al.\textsuperscript{24} stated that people with MCI, like healthy elderly people, were more likely to remember positively valenced stimuli while Fleming et al.\textsuperscript{1} found that mild-AD patients were more likely to remember negative information. There are therefore some inconsistencies in the literature. Findings appear to depend heavily on which tasks are used, the patient populations tested, and crucially, the stimuli and encoding instructions used.\textsuperscript{9} Emotional memory advantages tend to be more robust at delayed rather than at immediate recall and for recall rather than recognition tasks.\textsuperscript{25,26}

The 8- and 15-word test(s) from the Amsterdam Dementia Screening (ADS-6) battery\textsuperscript{27} are routinely used in neuropsychological assessment in Dutch hospitals and clinics to assess short-term episodic memory in early dementia. They are internationally known as the Rey Auditory Verbal Learning Test.\textsuperscript{28} Words are read out by the clinician one at a time and at the end of the first reading the client is invited to recall as many of the words as possible. The words are then read out again in a different order and the client is again asked to recall the words. The list is read out a total of 5 times and in this way the total number of words recalled across the 5 trials, the total number of repetitions, intrusions, and a learning curve can be obtained for each client. The data therefore give not only information on the total recall ability of a patient but also an indication of how they learn. A delayed recall and recognition score can also be obtained. In this study, however, we concentrated on immediate recall and learning.

The object of this study was to adapt the standard 8-word test to assess emotional memory in healthy elderly people, a MCI and in an AD population. Our goal was 2-fold. We wanted to develop lists that could easily be used by practitioners in the clinic and we wanted to determine whether our Dutch elderly population also showed an EMA. Most studies that have investigated memory for emotionally toned word lists have used very long lists and/or blocks. Elderly individuals find such tasks demanding and floor effects, especially in the demented population, are typical. Long lists are also not practical to use in everyday neuropsychological assessment when typically a number of tests are used to derive the overall cognitive functioning of a patient.

We expected to see better recall overall in the healthy control group compared to the other 2 groups with the MCI group showing a better memory overall than the AD group. We also expected to see EMAs in all 3 groups with recall better for the emotional than for the neutral words. Furthermore, a positive valence advantage for the controls and MCI groups\textsuperscript{8,24} and a negative valence advantage for the AD group was hypothesized.\textsuperscript{1} Finally, we had no specific expectancies with regard to the number of repetitions and intrusions the participants would make per list, nor did we predict how they would learn the lists. We know of no other study that has looked at learning in dementia using emotional word lists.
Materials and methods

Participants

This study included 95 participants who were either residents from 2 separate nursing homes in the Netherlands or recruited from the general population. They were divided into 3 groups—38 controls, 37 MCI, and 20 AD. Mini-Mental State Examination scores were obtained for all participants (MMSE). The controls had MMSE scores over 24 (mean = 27.4, SD = 1.7), had never received the diagnosis of dementia, lived in the community and had never suffered from a neurological or psychiatric disorder. They were also not taking psychiatric medications. These are also the criteria recommended by Petersen et al. for choosing control participants in dementia studies. Some researchers suggest a cutoff of 26 and on average our controls scored above this more lenient criterion thus making them comparable to other control groups found in the literature. The MCI group had MMSE scores ranging from 20 to 24 (mean = 21.5, SD = 2.0), lived in the community and had not (yet) received a diagnosis of dementia but who were believed to be functioning at a lower level (especially on memory) than the controls. The AD group had all received diagnoses of AD as defined by Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition (DSM-IV) criteria. These diagnoses were made on average 1 month before this study was carried out by a multidisciplinary team including a clinical neuropsychologist. The MMSE scores of the AD group ranged from 10 to 21 (mean = 16.2, SD = 3.9).

The groups were matched on age (controls: mean age = 81.5, SD = 6.4; MCI: mean age = 83.4, SD = 5.3; AD: mean age = 83.8, SD = 5.0), and on number of years of education (mean for all groups = 9 years). We could not however match the groups according to gender: there were more women in the MCI and AD groups (30 in the MCI and 17 in the AD) than in the control group (24). None of the participants were anxious or depressed (as assessed by the Hospital Anxiety and Depression Scale, HADS). Informed consent was obtained from all participants or their family members prior to participation.

Materials

Three new word lists, each 8 words in length, were developed using the Affective Norms for English Words list (ANEW) and then translated into Dutch. The words were carefully chosen according to their valence (pleasantness) level and were matched on all variables (including arousal/emotionality, imagability, and meaningfulness (as measured using the MRC Psycholinguistic Database)). Positive, neutral, and negative lists were therefore constructed. The 4 word lists and their descriptive statistics are shown in Table 1A and B.

The standard 8-word test was less arousing than the 3 newly constructed lists \(F(3,27) = 4.4, \: P < .025\). The neutral, positive, and negative word lists were matched on all variables (including arousal) and only differed on valence \(F(2,22) = 77.9, \: P < .001\).

Procedure

Participants were presented with the 3 new lists plus the standard 8-word test in 1 session with a 5-minute pause between each of the 4 lists. The instructions followed those provided with the standard 8-word test from the ADS-6 (see Introduction for a detailed explanation of this test), with each word list presented for 5 trials each, so that a testing session consisted of a total of 20 trials and 160 words. Participants were asked to recall as many words as they could at the end.
of each trial in any order that they remembered them. The order of the lists was counterbalanced across participants in each group in an attempt to control for possible order effects. Total number of correct responses per trial and overall for the 4 word lists were recorded as were the total number of repeats and intrusions for each of the lists per subject.

**Results**

Table 2 shows the mean total recall scores for the 3 groups over the 4 word lists plus the average “emotional score” (average across the positive and negative lists) and “non-emotional score” (total neutral list score). Also depicted in Table 2 are the recall means for each of the word lists and the mean HADS anxiety and depression scores per group.

Compared to the norms available for the standard 8-word test (derived from a healthy elderly population), controls scored at the 60th percentile, while the MCI group scored at the 20th percentile and the AD group scored below the 10th percentile. No norms (yet) exist for the new word lists.

An analysis of variance (ANOVA) was carried out for the total recall scores for the 3 groups collapsed over word lists and this revealed a main group effect \( [F(2,94) = 67.0, P < .001] \). Post hoc tests revealed that the controls recalled significantly more words over all word lists than both the MCI and the AD group and the MCI group recalled significantly more words than the AD group (\( P < .001 \) for all comparisons).

A second ANOVA was conducted to investigate whether there was an EMA. The average of the positive and negative lists was calculated per subject (positive total + negative total/2), and these values were compared to that person’s total recall score on the neutral list. In this way, we could compare an “emotional recall score” with a “non-emotional recall score” irrespective of the valence of the words (there is controversy in the literature as to whether a positive or negative advantage is seen in healthy and/or demented groups—see Introduction). An EMA was found \( [F(1,92) = 6.4, P < .025] \), with better recall for emotional words versus neutral words. This was the case for all 3 groups—there was no group x recall score interaction.
A third ANOVA of 3 (groups) by 4 (word lists—total scores) revealed a main effect of word list \[F(3,276) = 6.5, P < .001\] and post hoc tests revealed that participants performed best on the standard, positive, and negative word lists in comparison to the neutral list \(P < .05\) for all comparisons). No word list x group interaction was found.

A fourth ANOVA of 3 (groups) by 5 (trials) by 4 (word lists) was conducted. There was a main effect of word lists \[F(4,368) = 6.3, P < .001\] and post hoc analyses revealed identical results as depicted in the third ANOVA. There was also a main effect of trial \[F(4,368) = 191.9, P < .001\] and post hoc tests revealed that learning increased overall from trial 1 to 5. An interaction between trial and group was also found \[F(8,368) = 7.4, P < .001\] and this warranted separate ANOVAs per group. Controls showed increased learning across all 5 trials \[F(4,148) = 139.9, P < .001\]. The MCI group also showed steady learning across the 5 trials \[F(4,144) = 77.3, P < .001\], and there was also a main effect of word list for this group \[F(3,108) = 5.9, P < .003\]. This reflected better recall over all 5 trials for the negative list than for the neutral list \(P < .01\). The AD group also showed evidence of learning \[F(4,76) = 27.0, P < .001\] despite the fact that their learning curves were flatter and more variable than those of the other 2 groups. No other significant effects emerged. Learning curves were plotted (see Figure 1) to compare the 3 groups’ recall for all 4 word lists.

A further ANOVA was carried out to assess whether there were any list order effects. No main effect of list order or interactions of list order x word list or list order x group emerged suggesting that our counterbalancing was effective and that list order had no effect on recall performance. Finally, ANOVAs were carried out to assess list order effects, and repetitions and intrusions per word list in all 3 groups. Only the significant findings are reported here. A main group effect was found for intrusions \[F(2,94) = 3.8, P < .03\] with more intrusions made by the AD group in comparison to the controls \(P = .06\) and MCI \(P = .04\) groups and no difference evident between controls and MCI. There was also a main list effect for intrusions \[F(3,276) = 8.14, P < .001\] with more intrusions made over all groups for the neutral \(P = .02\) lists compared to the standard, positive, and negative lists.

**Discussion**

Several interesting findings emerged from this study. Controls had, unsurprisingly, better recall overall than the other 2 groups and the MCI group had better recall overall than the AD group. This result reflects the fact, especially if we consider the MCI group to be in the very early stages of AD, that episodic memory declines as AD progresses. We also found an EMA with better memory for emotional than nonemotional words. Interestingly, this was the case for all 3 groups and could not be said to be due to any differences in anxiety or depression (see Table 2 for HADS scores). Although Fleming et al did not directly measure mood, these authors indicated that this variable may have an important bearing on how well participants recall emotional stimuli (depressed people being more likely to recall negative stimuli). None of our participants were clinically anxious or depressed so their better recall of emotional stimuli cannot be attributed to their mood. Our findings are partially consistent with the work of Brierly et al and Fleming et al and suggest that emotional stimuli may help not only healthy elderly people but also people with MCI and AD to remember better. We did not however find any recall differences between the negative and positive lists for any group and therefore our results do not support the positivity effect (see Introduction) found by some researchers in elderly individuals.

The learning curves suggested that all 3 groups showed evidence of learning over the 5 trials. For controls, this was the case for all 4 word lists. For the MCI group better recall occurred for the negative list than for the neutral list over the 5 learning trials, a finding which supports the negative bias found for Alzheimer patients by Fleming et al. These authors’ patient population had a mean MMSE score of 21 making them quite similar to our MCI group. The AD group in this study also showed evidence of learning while their curves were flatter and more variable than the other 2 groups and their recall appeared to peak between trials 3 and 4 and drop off again by trial 5 for the 3 new word lists (for the standard word list, they showed a peak at trial 3). This suggests that even this group can recall words in an episodic memory test albeit at a lower level than controls or those with MCI. Severely reduced episodic memory is a hallmark of AD and therefore this result is not overtly surprising. Our results do suggest however that all 5 trials may not be necessary to test episodic memory in a moderately demented AD population, 4 trials may be adequate, especially if we are simply interested in the best performance these patients can achieve.

Interestingly, Brand and Jolles suggested that flatter learning curves obtained in a serial list
learning task similar to the one used in this study could be due to frontal neocortical dysfunction and this would be reflected by more repetition errors. Our AD group showed (nonsignificant) signs of more repetition errors and (significantly) more intrusion errors compared to the other 2 groups. Although hippocampal atrophy is generally the first sign of neuronal damage in AD, it is possible that some patients in our AD group also had frontal lobe damage. Unfortunately, we do not have magnetic resonance imagings (MRIs) which could confirm or dispute this suggestion. Obtaining brain scans is therefore something to consider for future research.

Our study also revealed that the standard 8-word test is not a “neutral” test as many practitioners believe but rather a positively valenced one (see Table 1B). It is therefore appropriate to use in the clinic when the neuropsychologist wants to obtain the best performance possible from the patient.

Figure 1. The learning curves of the 3 groups for each word list.
There are several limitations present in our study. The MMSE is a relatively crude cognitive screen which cannot tell us very much about the specific cognitive functioning of our participants. We would recommend more detailed assessment in future studies, for example, using the Cambridge Cognitive Examination (CAMCOG). This instrument can provide much more information on the cognitive abilities of the participants and would have helped us to determine whether the MCI group was actually suffering from mild AD. Nothing however can match a clinical diagnosis and we emphasize the continuing need for such a diagnosis in future studies.

Another limitation is that we could not adequately assess possible gender effects—there were more women than men in our population. It is difficult to see how this could be prevented in future studies because women do tend to live longer than men and our population was made up of old-elderly individuals thus perhaps exacerbating the preponderance of women. This might not be such a problem however because Kensinger et al failed to find gender effects in their large demented population. Interestingly, EMAs were not found for the AD patients (regardless of disease severity) in Kensinger et al’s study, while we found EMAs for all 3 groups tested here. As suggested in the Introduction, encoding instructions, stimulus characteristics and populations studied may all determine whether EMAs are found or not. Furthermore, we only used a recall task because these types of tasks tend to produce EMAs more readily than recognition tasks. It would however also be useful to include recognition tasks in future studies.

Charles et al suggested that in normal aging people would shift from a better memory for negative information (found in young individuals) to a better memory for positive information. This study does not support these findings—we found better memory for both negative and positive words (versus neutral) in all groups. These finding have implications for treatment. More successful cognitive rehabilitation strategies could feasibly be designed, for example, by focusing on positive feedback and using emotionally toned words and phrases throughout treatment. Future studies could assess the usefulness of this idea by manipulating instructions during treatment and measuring its effectiveness on both short- and long-term cognitive and behavioral functioning.

In conclusion, although we were successful in creating emotional word lists which could be used by practitioners in the clinic as part of a standard neuropsychological assessment, we recognize that these lists will need to be validated and norms obtained before they can be widely used by clinicians. It is often difficult to find suitable emotional memory tests to use with patient groups because, as we suggested in the Introduction, lists typically used in laboratories are often too long and complex to be used with elderly people and demented patients. Our lists provide a possible solution to this dilemma. Also, despite the fact that the average age of our population (~80 years for all 3 groups) was older than is found in most of the literature, our participants performed relatively well on the tests. This gives not only elderly people hope (their memory may not be completely damaged with advancing age) but also suggests that memory research can be conducted even in the oldest individuals. Finally, the present findings suggest that using emotional stimuli may help improve memory recall in the very old even if the participants have been diagnosed with AD. Future research will determine whether this EMA holds for other kinds of stimuli in these populations and will assess how they learn and determine whether everyday memory may also be enhanced via manipulating the emotional connotations of the to-be-remembered events.

References

6. Libon KJ, Bogdanoff B, Cloud BS, et al. Declarative and procedural learning, quantitative measures of the...


